Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
AIR FORCE PERSONNEL AND TRAINING RESEARCH CENTER
P. O. BOX 1537, LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

DAVID A. GRANT, Associate Editor
UNIVERSITY OF WISCONSIN

DELOS D. WICKENS, Associate Editor
OHIO STATE UNIVERSITY

CONSULTING EDITORS

E. James Archer, *University of Wisconsin*
Judson S. Brown, *State University of Iowa*
Cletus J. Burke, *Indiana University*
Paul M. Fitts, *Ohio State University*
Frederick C. Frick, *Massachusetts Institute of Technology*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
Harold W. Hake, *University of Illinois*
Lloyd G. Humphreys, *AF Personnel and Training Research Center*
Arthur L. Irion, *Tulane University*
Howard H. Kendler, *New York University*
Donald B. Lindsley, *University of California, Los Angeles*
Kenneth MacCorquodale, *University of Minnesota*
Quinn McNemar, *Stanford University*
Neal E. Miller, *Yale University*
Kenneth W. Spence, *State University of Iowa*

Arthur C. Hoffman, Managing Editor
Helen S. Orr, Assistant Managing Editor
Editorial Staff: Genevieve A. Highland, Frances Hazel, Sadie J. Doyle

CONTENTS

Editorial Announcement 1

The Role of Redundancy in the Discrimination of Visual Forms: MAURICE RAPPAPORT 3

Stimulus Recall Following Paired-Associate Learning: SAMUEL M. FELDMAN AND BENTON J. UNDERWOOD 11

The Role of Meaningfulness (*m*) in Paired-Associate Verbal Learning: CLYDE E. NOBLE AND DELDON A. MCNEELY 16

A Re-Examination of Rubin's Figural Aftereffect: IRVIN ROCK AND IRWIN KREMEN 23

Experiments on Figural Dominance: THEODOR M. KÜNNAPAS 31

Reinforcement of Verbal Behavior: DOUGLAS M. MCNAIR 40

The Effort Variable in the Acquisition, Extinction, and Spontaneous Recovery of an Instrumental Response: EDWIN G. AIKEN 47

Self-Pacing Performance as a Function of Perceptual Load: R. CONRAD AND B. A. HILLE 52

Stimulus Conditions and Retroactive Inhibition: JOEL GREENSPOON AND REDGE RANYARD .. 55

Retroactive Inhibition as a Function of the Degree of Original and Interpolated Learning: GEORGE E. BRIGGS 60

The Relation of Decision-Time to Stimulus Similarity: WILLIAM N. DEMBER 68

American Psychological Association

Vol. 53 No. 1 January 1957

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $8.00, or $16.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal, experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research effort is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this JOURNAL, 1957, 53, 1–2).

Address all articles submitted for publication to the editor, Arthur W. Melton, Air Force Personnel and Training Research Center, P. O. Box 1557, Lackland Air Force Base, San Antonio, Texas. All manuscripts should be submitted in two complete copies; only the original copy will be returned if the article is not accepted for publication. Figures may be original line drawings or glossy photographic prints, but should not exceed 8½ by 11 in. Care should be taken to insure legibility of labels and critical features of drawings when reduced to a single-column or (exceptionally) a double-column figure in the JOURNAL.

To save printing costs supplementary material to articles in this Journal are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 [4, Suppl.], 389–449). In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Se*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879. Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1957 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 53, NO. 1 JANUARY, 1957

EDITORIAL ANNOUNCEMENT

Henceforward this JOURNAL will publish brief reports of simple extensions of previously reported findings and replications of previously reported experiments by a special arrangement which provides the opportunity of relatively rapid publication but at minimal cost in space in the JOURNAL. Such reports will be limited to not more than one and one-half printed pages, with not more than one page as the optimal length. Reports of one page or less will be published whenever, in the make-up of an issue, there would have been a blank page at the end of the issue; reports of one and one-half pages in length will be published whenever the last page of an issue would have been blank and the preceding page was half blank. Within the terms of this arrangement, these reports will be published in the order of date of receipt and should therefore provide earlier publication as well as more complete utilization of the pages of the JOURNAL.

Reports that are to receive this special treatment must include in their titles the term "Supplementary Report: . . ." or "Replication Report: . . . ." By supplementary report is meant a simple extension of a previously reported experiment to new values of a constant or variable factor previously employed. For this reason, it is expected that the report will be written with the assumption that the reader has the previous report before him. Therefore, details of the previous introduction, method, results, and discussion need not be repeated and the article should summarize the additional results in the space allowed. If the extension of the previous work should lead to unexpected findings, or should for any other reason be given more extensive treatment, the article should be submitted as a regular JOURNAL article. If the Editor disagrees and believes the more extensive report should be reduced to a "Supplementary Report," the author will be given an opportunity to reduce his article to the required length without change in the "date of receipt" and its privileges.

In the case of "Replication Reports," the same principles as described above will apply. The only difference is that the report is a repetition of a previously reported experiment rather than a supplement to it. Since replication is rarely exact, there will be the question whether the report concerns a repetition of previous work or a supplement to it. Most often this question can be resolved by the intent of the investigator, as shown in his control and manipulation of factors in the experiment. If, analogous to the unexpected supplementary find-

ing, a replication report contains data significantly at variance with those previously reported, the report may be submitted as a regular JOURNAL article and treated accordingly.

The key to the preparation of a supplementary or replication report is the assumption that it will be read in the context of the previous published report. Measures of the dependent variable should be common to the original and supplementary report or replication report and should require no new explanation. The data reported should have the same tabular or graphic form as that used in the original article, in order to facilitate comparisons. If figural presentation of data is used, it may be possible to repeat some of the data of the previous study without additional cost in space, but tabular data from the previous study should not be reprinted. Figures should be planned for one-column presentation, if used at all. References should most often consist only of the reference to the previous article, and the introduction and discussion sections should be held to a few sentences.

It is hoped that the innovations described above will provide an avenue of publication for studies that hitherto have gone unreported. It is also hoped that they will provide an alternative to outright rejection of articles of greater length in which the new contribution has not seemed commensurate with the cost in publication space.

Whether these hopes are realized will depend on the reality of these apparent needs and on the success of contributors in achieving brief but meaningful reports.

ARTHUR W. MELTON
Editor

Journal of Experimental Psychology
Vol. 53, No. 1, 1957

THE ROLE OF REDUNDANCY IN THE DISCRIMINATION OF VISUAL FORMS[1]

MAURICE RAPPAPORT[2]

The Ohio State University

Redundancy has been shown to be an important factor in written language communication (**6, 8, 9**) and in communication engineering (**4**). In addition, evidence has been accumulating to indicate that redundancy may also be important in form perception. Gestalt psychologists have emphasized that symmetrical figures, which may be considered redundant, are more readily remembered and identified than are asymmetrical figures (**5**). It has been demonstrated that redundant features of a figure contribute to the ease with which *S* can guess the form that an initially unknown stimulus will eventually take (**2**). Stimulus redundancy, however, does not always help discrimination. The type of redundancy used may alter considerably the ease with which visual forms are identified (**3**).

Redundancy may be introduced by encoding a message so that there are more symbols or stimulus features than the theoretical minimum necessary to transmit a particular amount of information. In this way redundancy, defined as one minus the ratio of actual to potential information, may provide an effective means for overcoming perturbation in a message (stimulus) brought on by noise in a communication channel (**10**). A noise source, as the term is used in reference to a communication channel, has been defined as any source that produces signals which are not coherent with any signals to which meaning is assigned (**4**). Examples of different methods of varying the redundancy of visual patterns, and the levels of noise affecting these patterns, have been given in a previous report (**3**).

Noise affecting visual perception may be conceived as arising from either internal or external sources. Noise of internal origin may conceivably be associated with many different physiological or psychological variables, while external noise may be associated with factors affecting the generation and/or presentation of the stimulus to the eye.

Four experiments were carried out in order to investigate the role of redundancy in visual perception. The effects of two types and five amounts of visual redundancy introduced under two conditions of visual noise were studied, with discrimination time as the dependent variable. The first two experiments were designed to test the hypothesis that there is an inverse relation between discrimination time and the amount of redundancy in a visual stimulus observed in a relatively noise-free situation.

Experiments 1 and 2

Method

Stimuli.—The stimulus material consisted of constrained metric figures whose general method

[1] This report is based on a dissertation submitted to the Graduate School of The Ohio State University in partial fulfillment of the requirements for the Ph.D. degree in psychology. The research was supported by Contract Nonr-495(02) with the Office of Naval Research and was conducted in the OSU Laboratory of Aviation Psychology. The writer is greatly indebted to Professor P. M. Fitts under whose direction this investigation was conducted.

[2] Now Associate Research Psychologist at Stanford Research Institute, Menlo Park, California.

of construction has been reported previously (**3**). Samples of eight figures were used in each experiment. They were selected at random from a parent population of 24 figures. The parent population consisted of all solid asymmetrical figures that can be constructed in a 4 × 4 matrix by drawing four bars such that a bar is one of four lengths and all bars are of a different length. Adopting information notation, the presentation of one of these 24 constrained asymmetrical figures can convey a maximum of $\log_2 24$ or 4.58 bits of information on the average. Since the maximum information that can be conveyed by a random asymmetrical figure constructed in the same matrix, i.e., in which each bar could be any one of four lengths, is $\log_2 256$ or 8 bits, the relative redundancy of each of the 24 constrained asymmetrical figures is approximately 43%. Making each figure symmetrical brings the relative redundancy value to 71%.

Four further steps of redundancy were introduced by repeating the figure contour a different number of times so that figures with 81, 86, 91, and 93% redundancy were also obtained. In order to hold over-all area constant, bar width was decreased proportionately with each increase in the number of contour details. (The effects of decreasing size of detail were investigated in Exp. 3.)

Two different methods of introducing redundancy were employed. The result will be called congruent redundancy and reciprocal redundancy. Congruent redundancy consists of a simple repetition of a basic four-unit contour pattern; thus, if the contour of a figure consisted of bars 4, 1, 3, and 2 units of length, then congruent redundancy is introduced by repeating the pattern in the order 413241324132 . . . etc. Reciprocal redundancy consists of the repetition of bar lengths so that each preceding four-bar pattern is repeated in reversed order; thus, if a contour pattern consists of bar lengths of the order 4132, then reciprocal redundancy is introduced by repeating the pattern in the order 413223144132 . . . etc.

Two of the figures used in Exp. 1, each displaying two types and the five different amounts of redundancy, are shown in Fig. 1. In each column the same basic symmetrical pattern is made more redundant in the five steps. Similar figures randomly selected from the same parent population of 24 figures were used in Exp. 2.

Subjects.—Twenty-one college students (15 male, 6 female) at the junior level or above served as *Ss*. They were paid for their time.

Procedure.—In Exp. 1 all *Ss* were tested under 10 conditions of a standard recognition task (**3**). In five of the conditions reciprocal mirror-image figures with 71, 81, 86, 91, and 93% redundancy were employed and in the remaining conditions congruent figures with the same five steps of redundancy were used. Each *S* served for five sessions. They were given two conditions per session. Each condition was repeated three consecutive times on the day it was given. The sequence of conditions, and the order of working with individual stimuli within conditions, were randomized for each *S*.

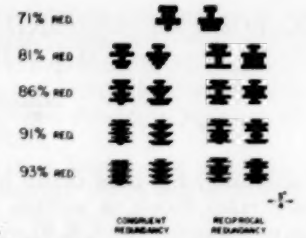

FIG. 1. Congruent and reciprocal redundancy. Each column shows the same stimulus figures made 81, 86, 91, and 93% redundant. The basic symmetrical figure at the top is 71% redundant.

The recognition task was administered in a manner essentially the same as described previously (**3**). The actual task of *S* was to sort out a specified pattern from among eight similar patterns. The *S* was required to make this discrimination six times per figure for a total of 48 discriminations for each condition. Scores were averaged over the last two of the three repetitions of each condition. Average recognition times rather than error scores were used as an index of the effect of redundancy on the discrimination of visual forms since errors were negligible due both to the nature of the recognition task and to the instructions given *S*.

Analysis of the data for Exp. 1 indicated that three trials on the same condition permitted *S* to reduce the absolute time needed to sort the stimuli but did not change the relations among scores for different conditions. In the light of these results it appeared feasible to administer all 10 conditions in a single session, thus controlling quotidian variability. Experiment 2 consisted of a replication of Exp. 1 with this modification of experimental procedure.

Results

The effect of visual redundancy on sorting time using the two random sets of figures is shown in Fig. 2. Each point on the graph represents the mean sorting time in seconds required by the 21 *Ss* in Exp. 1 and

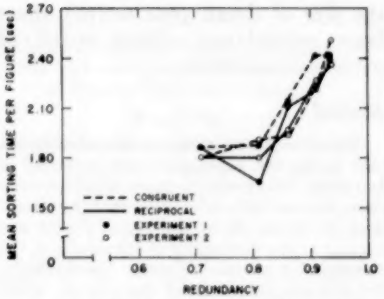

FIG. 2. Mean time in seconds for sorting redundant figures in Exp. 1 and Exp. 2.

the 18 *Ss* in Exp. 2 to identify an average figure under each of the 10 experimental conditions. The results in the two experiments are similar. In general, an increase in sorting time accompanied an increase in redundancy when either reciprocal or congruent patterns were presented free of external perturbation. This effect was most apparent for figures that were more than 81% redundant.

The results of the analysis of variance for each experiment are presented in Table 1. Both analyses fail to warrant rejection of the null hypotheses that the two types of redundancy are equivalent. The effect of level of redundancy, however, is significant at $P < .01$ in both experiments.

The differences between successive pairs of treatment means representing the effects of the use of successive levels of visual redundancy were evaluated by individual t tests. The resulting t's in both experiments were all significant at the $P < .001$ level except for the differences between 71 and 81% redundancy for congruent figures in Exp. 1 and for reciprocal figures in Exp. 2. These differences were not significant.

An examination of the recognition times for individual figures did not reveal any one figure which was markedly easy or difficult to discriminate in Exp. 1. In Exp. 2, however, a "staircase" figure whose contour details happened to fall in the order 1234 units was sorted consistently faster in 9 out of 10 experimental conditions, and was sorted second fastest on the remaining condition.

Discussion

The results of both experiments lead to the tentative conclusion that in a relatively noise-free situation redundancy has a decidedly retarding effect on speed of discrimination. Thus, the hypothesis

TABLE 1

ANALYSIS OF VARIANCE OF SORTING TIMES FOR TWO TYPES AND FIVE AMOUNTS OF REDUNDANCY IN EXP. 1 AND 2

Source	Experiment 1			Experiment 2		
	df	Mean Square	*F*	*df*	Mean Square	*F*
Type of Redundancy (T)	1	.4601	3.812	1	.0137	0.266
Level of Redundancy (R)	4	3.1098	19.720**	4	2.7281	41.024**
Subjects (S)	20	1.7676		17	2.3457	
$T \times R$	4	.1241	.6144	4	.1045	1.869
$T \times S$	20	.1207	.598	17	.0515	0.921
$R \times S$	80	.1577	.780	68	.0665	1.189
$T \times R \times S$	80	.2020		68	.0559	

** $P < .01$.

that decreased identification time is associated with increasing redundancy in visual figures is not supported.

In the course of the two experiments certain points arose which made it necessary to consider other factors that might be more closely associated with the results than the factor of redundancy. One point stemmed from *Ss'* verbal reports. Almost unanimously they agreed that once they became familiar with the sorting task they made little or no effort to look at the whole figure but examined instead some part of the figure that was sufficient to identify it. If one detail was not sufficient for positive identification (there may have been one or two bars of the same length and in the same positions in another figure), then *Ss* reported that they examined two or three details. These introspective reports suggested that the redundant aspects of a figure were not being utilized.

A second point, which must be considered in interpreting the present data, is detail size. Over-all area was held constant in all of the figures. Introducing more and more contour details (bars) while holding over-all figure area constant required that the width of the bars be reduced proportionally with each increase in the level of redundancy. Thus, as redundancy was increased *Ss* were required to discriminate smaller and smaller details. The effects of size of detail are of special importance particularly if it is true that *S* does not survey the characteristic contour of the whole pattern but, instead, concentrates only on some distinctive part of the figure.

These considerations suggested that a further study be made where the effects of size of detail could be more closely examined.

Experiment 3

In the two previous experiments the introduction of redundancy was accompanied by the reduction of the size of details and hence the effects of these two variables were confounded. The purpose of the present experiment was to determine whether the size of detail (bar width) apart from redundancy affects speed of pattern recognition.

Method

Stimuli.—The eight basic symmetrical figures used in the first experiment were presented in four sizes. The reduction in size took place only about the vertical axis in order that the dimensions of the details in the figures should correspond to the dimensions of the details in the corresponding redundant figures. Consequently, the over-all dimensions of the area in which the figures were presented were .75 × .50 in., .75 × .38 in., and .75 × .19 in. These figures were all 71% redundant, but the size of the contour details drawn in these areas correspond to the size of details in the previously used figures with redundancy values of, respectively, 81, 86, and 93%. Both types of figures are illustrated in Fig. 3. The top figure represents the basic constrained symmetrical metric figure common to both groups, i.e., the original contour pattern before either a change in redundancy or a change in size of detail was introduced.

Subjects.—Sixteen of the *Ss* that had been used in Exp. 2 served in Exp. 3. The method was the same as that employed in the previous experiments.

Results and Discussion

The mean sorting times and *SD*'s in seconds for both the figures with different amounts of redundancy and the figures with only 71% redundancy but matched for size of detail with the more redundant figures are shown in Table 2. It will be seen that an increase in sorting time accompanied

Fig. 3. Stimulus figures matched for size of details. The figures on the right differ from those on the left in area and amount of redundancy. At the top is the basic figure before either increases in redundancy or decreases in size of detail have been introduced.

TABLE 2

SORTING TIME FOR STIMULI MATCHED FOR SIZE OF DETAIL BUT DIFFERING IN REDUNDANCY AND IN AREA

(N = 16 Ss)

	.125-In. Detail		.094-In. Detail		.047-In. Detail	
	81% R	71% R	86% R	71% R	93% R	71% R
Mean	1.57	1.62	1.86	1.82	2.26	2.17
SD	.321	.277	.351	.283	.419	.460

each increase in redundancy and also each decrease in size of detail. None of the three t tests between pairs of figures having the same sizes of contour details was significant. There is no reason, therefore, to reject the null hypothesis that the effects of redundancy on identification time can be accounted for by the effects of size of figure details when redundancy is introduced by repeating some part of the figure contour while holding overall area constant.

The results of Exp. 1 and 2 may be largely accounted for by the effects of reduced detail size rather than the amount of repetition of the contour present in any one figure. The fourth experiment in this series was carried out in order to determine whether the empirically determined detrimental effects of small size of detail or the theoretically beneficial effects of redundancy has the greater effect on speed of recognition of visual patterns under conditions of external perturbation.

EXPERIMENT 4

It is proposed that an increase in the redundancy of a visual stimulus pattern can facilitate its identification under unfavorable viewing conditions. This proposition is based on the assumption that in the noise-free case S does not utilize the entire configuration of a visual stimulus when he identifies it but uses, instead, certain critical parts of the figure as important discriminatory cues. On these grounds the further assumption is made that, within limits, the more repetitive the figure contour, i.e., the more probable it is that all of these cues will not be changed by noise, the more likely it will be that S will be able to identify a stimulus correctly.

In Exp. 4 the effects of redundancy on recognition time are investigated when the metric figures are distorted by background visual noise. The hypothesis to be tested is that sorting time is inversely related to the amount of redundancy in a visual stimulus when the stimulus is distorted by this type of visual noise.

Method

Stimuli.—The stimuli employed in Exp. 4 were the same eight congruent redundant metric figures used in Exp. 1. The basic asymmetrical figure was added in order to increase the range of redundant values, and the 81 and 91% redundant figures were omitted since they represented intermediate values that were not critically related to the hypothesis.

At each of the four different levels of redundancy, figures were perturbed by 16.67% background noise. This level of noise was chosen on arbitrary grounds because it appeared to distort a figure sufficiently to cause difficulty of recognition and yet not to distort it to the extent that identification was impossible.

The procedure for obtaining patterns of background noise was as follows. The white cells constituting the background of each figure were numbered. One sixth of these white cells were selected randomly and were made black like the figure. For figures with redundancy values of 43, 71, 86, and 93% the total number of background cells affected were 1, 2, 4, and 8, respectively. It should be noted that the method used for generating background noise permitted an a priori quantitative graduation of levels of perturbation. Earlier research involving graded distortions of visual patterns had usually to rely on *post hoc* ranking by judges, as for instance in the method introduced by Yum (11). Some typical perturbed figures are shown in Fig. 4.

The eight figures for each level of redundancy

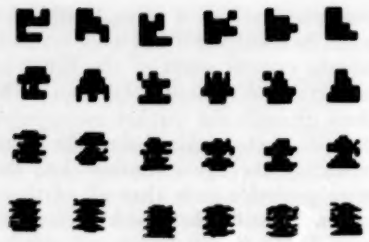

FIG. 4. Redundant figures perturbed by approximately 17% visual noise. Each column shows the same stimulus figure made 43, 71, 86, and 93% redundant.

were perturbed by three different random samples of noise. Each noise pattern appeared twice among the 48 figures visible on a sorting board. Otherwise the sorting board was arranged as described previously (3). The test figure, which was shown at the top of the sorting board, was free of noise. Error as well as time scores were recorded.

Results

The results of this experiment are shown in Fig. 5. When figures were distorted by background noise there was a marked increase in the time required to identify them. However, on the average, a decrease in sorting time was associated with each increase in redundancy.

Table 3 presents the results of the analysis of variance of the data from this experiment. The effect of

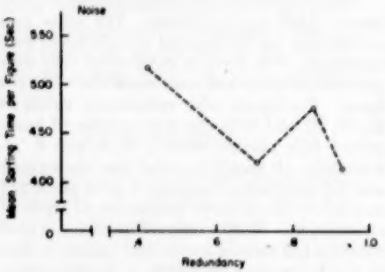

FIG. 5. Mean time in seconds for sorting redundant figures that have been perturbed by noise.

TABLE 3

ANALYSIS OF VARIANCE IN SORTING TIME FOR EXP. 4

Source	*df*	*MS*	*F*
Level of Redundancy (*R*)	3	21.18	37.821**
Subjects (*S*)	15	.73	1.304
R × *S*	45	.56	

** $P < .01$.

amount of redundancy on sorting time is significant at the .01 level. The results, therefore, support the hypothesis that recognition time is inversely related to the amount of redundancy in a visual stimulus pattern that has been distorted by background noise. When the differences between pairs of treatment means were tested, it was determined that it took significantly less time to sort figures with 71, 86, and 93% redundancy than it did to sort figures with only 43% redundancy (all *t*'s, $P < .01$). This significant effect occurred despite the decreases in size of details associated with each increase in redundancy.

The time taken to sort figures with 86% redundancy was significantly greater than the time taken to sort figures with either 71 or 93% redundancy (in both cases $P < .01$).

As was expected, the number of errors in identifying different patterns under each experimental condition corresponded closely to their average sorting times. The mean number of errors for figures with 43, 71, 86, and 93% redundancy were, respectively, 10.13, 8.88, 9.44, and 7.44, whereas the mean times were 5.15, 4.21, 4.76, and 4.15 sec., respectively.

DISCUSSION

The relatively slow sorting time in Exp. 4 for figures with 86% redundancy

is inconsistent with the other obtained scores. An examination of the three random noise samples did not reveal a spurious effect attributable to any one of them. A rise in both sorting time and error scores accompanied all three samples. This suggests that there may be an interaction between certain amounts and types of noise and certain amounts and types of redundancy. The hypothesis of interaction cannot, of course, be tested with the present data. Further research might profitably be directed, however, to investigations on the effects of ratio of noise size to the size of stimulus details and also to the manner in which noise is introduced (see 3). For example, background noise which can only add to a visual pattern, as in the present experiment, may have a different effect on discrimination than noise which can both subtract and add to a contour of a visual figure.

Some of the concepts that have been used in these experiments stem from information theory. These concepts make possible a rigorous approach to some of the problems of visual perception. They may also be helpful in the re-evaluation of certain older ideas that have arisen in the study of particular problems of form discrimination. For example, the Gestalt notion of "good" figure emphasizes, among other things, symmetry and simplicity. The present experiment demonstrates how symmetry may be systematically and quantitatively investigated in terms of the redundant aspects of a figure contour.

Simplicity as a stimulus correlate of form perception may be investigated from the point of view that the simplest figures come from a more limited population of stimuli that *S* has learned to expect under a given set of circumstances. In this respect it would be of value to assess the effects on form discrimination of the knowledge *S* has about (*a*) the subset of stimuli that will be present in the field of view at any one time (e.g., *S* would be led to expect a particular set of eight figures), (*b*) a small population of stimuli from among which are drawn the sample that may appear before him in a particular situation (e.g., *S* would be led to expect eight figures out of a possible sixteen figures), or (*c*) a very large population of stimuli from among which a particular sample might appear (e.g., *S* would be led to expect any figure whatever).

Summary

Four experiments were performed to determine the effect of two types and five levels of redundancy on *S*s' ability to recognize visua patterns under two conditions of visual noise. When background noise was present redundancy was found to facilitate rapid discrimination; when stimulus noise was absent an increase in redundancy was associated with an increase in recognition times. The longer times in the noise-free case appeared to be attributable primarily to the difficulty of discriminating smaller details in the figures which were more redundant.

No differences in sorting times were found for figures with two different types of redundancy presented in the noise-free situation. The lack of differences was interpreted as due, in part, to the nonutilization of redundant features of a stimulus by *S* when external noise was absent.

It was suggested that in certain situations there may exist a balance of effects between the beneficial characteristics of various types of redundancy and the detrimental features that arise when redundancy is introduced in too great amounts or in an ineffective way. Implications of the present approach for an interpretation of two traditional Gestalt figural concepts were mentioned.

References

1. Adams, O. S., Fitts, P. M., Rappaport, M., & Weinstein, M. Relations among some measures of pattern discriminability. *J. exp. Psychol.*, 1954, **48**, 81–88.
2. Attneave, F. Some informational aspects of visual perception. *Psychol. Rev.*, 1954, **61**, 183–193.
3. Fitts, P. M., Rappaport, M., Weinstein, M., Anderson, N., & Leonard, J. A. Stimulus correlates of visual pattern recognition: a probability approach. *J. exp. Psychol.*, 1956, **51**, 1–11.
4. Goldman, S. *Information theory*. New York: Prentice-Hall, 1953.

5. Koffka, K. *Principles of Gestalt psychology.* New York: Harcourt, Brace, 1935.
6. Miller, G. A. *Language and communication.* New York: McGraw-Hill, 1951.
7. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* New York: Houghton, Mifflin, 1953.
8. Newman, E. B., & Gerstman, L. S. A new method for analyzing printed English. *J. exp. Psychol.*, 1952, **44**, 114–125.
9. Shannon, C. E. Prediction and entropy of printed English. *Bell Sys. Tech. J.*, 1951, **30**, 50–64.
10. Shannon, C. E., & Weaver, W. *The mathematical theory of communication.* Urbana: Univer. Ill. Press, 1949.
11. Yum, N. S. An experimental test of the law of assimilation. *J. exp. Psychol.*, 1931, **14**, 68–82.

(Received December 19, 1955)

Journal of Experimental Psychology
Vol. 53, No. 1, 1957

STIMULUS RECALL FOLLOWING PAIRED-ASSOCIATE LEARNING

SAMUEL M. FELDMAN AND BENTON J. UNDERWOOD

Northwestern University

In the present experiment *S* learned a list of verbal paired associates by standard procedures. Immediately after learning he was presented with the response words, one at a time, and was asked to give the stimulus word with which that response had been assocated during learning. Thus, if during learning *S* was required to anticipate B when A was presented, on the recall he was asked to give A when B was shown. So far as is known, results of such operations have not been reported previously. The basic question, therefore, is whether or not *S* learns the R-S association when instructed merely to learn the S-R association. Thornton performed a study (unpublished) in the Northwestern University laboratory which answers this basic question. A list of 10 paired adjectives was learned to a criterion of two consecutive perfect trials. Immediately afterward *S* was shown each response word separately and was asked to give the stimulus word. On the average 83% of the stimuli were correctly recalled. This finding clearly suggested that *S* does learn R-S associations while learning S-R associations and learns them rather well. The present study will, first of all, endeavor to confirm Thornton's findings and will, secondly, explore the influence of stimulus and response similarity on this recall phenomenon.

The phenomenon in question may be associated with what is called, commonly, *incidental learning*. In incidental learning (e.g., **2**) *S* learns material when he is not specifically instructed to do so. In the present situation *S* appears to learn associations which he is not instructed to learn but during a period when he is learning other associations which he has been instructed to learn. The fact that he is instructed to learn something differentiates the operations from those commonly used to study incidental learning. The high recall shown by Thornton's *S*s might argue against incidental learning and suggest something more fundamental. It might, for example, suggest that in learning S-R connections the acquisition of R-S associations takes place as an integral part of the learning. Ideas of direction of association (S → R) which are used to conceptualize experimental operations may not at all reflect the learning which is taking place. Nevertheless, in initiating the present study R-S learning was viewed as essentially a form of incidental learning as a basis for exploring the influence of variables of which it might be a function.

In order to minimize associative interactions between stimulus and response words, nonsense syllables were used as stimuli and adjectives as responses. In one case high intralist stimulus similarity was used, and in another low. It was believed that with high intralist similarity *S* would be forced to attend more carefully to the stimuli than when similarity was low, thus facilitating R-S learning. Response similarity was also high in one list and low in another. If this high-response similarity forced *S* to pay careful attention to the responses,

R-S learning might be inhibited as compared to the case of low-response similarity.

Method

Materials and subjects.—Four lists of seven paired associates each were constructed. The stimuli were nonsense syllables taken from Glaze (1) with an average association value of 11.4%. The adjectives used as responses came from a previous study (5). The lists are shown in Table 1. For high similarity among stimuli, two consonants were repeated four times and two others were repeated three times each. For low similarity, no consonants were repeated. In the case of responses, high similarity was produced by having a commonality of meaning among all seven words; for low similarity, words were used which had little commonality of meaning.

A total of 120 undergraduate students served as *S*s. Thirty *S*s were assigned to each of four groups, the groups being differentiated only by the list which they learned. These groups will be referred to by the symbols given above the lists in Table 1; e.g., HS-HR, indicates the list had high stimulus similarity and high response similarity. The 24 permutations of the order of the four lists were used as a basis for assigning *S*s to a given list. An *S* who failed to learn the list assigned was replaced immediately by the next *S* coming to the laboratory. One *S* assigned to the HS-HR group failed to learn as did one assigned to LS-LR group.

Procedure.—Lists were presented on a Hull-type drum at a 2:2-sec. rate, i.e., 2 sec. for the stimulus and 2 sec. for the stimulus and response together. Three different orders were used to minimize serial learning. Learning was taken to two successive perfect trials. Immediately after attaining the criterion, *S* was presented the responses (adjectives) at a 4-sec. rate (the adjective was seen for only the first 2 sec. of the 4-sec. period) and was asked to give the stimulus (syllable) which went with each adjective. During this stimulus recall *S* spelled the syllables. The *S* never saw the original stimuli during this recall period. Following this drum recall, *S* was presented with the list of adjectives on a sheet of paper and again was asked to call out the syllables (stimuli) for each. No time limit was imposed.

Results and Discussion

Learning.—The mean number of trials to learn each list, and the mean number of errors per trial in learning are shown in Table 2. Looking first at trials to learn, performance varies significantly as a function of both stimulus similarity and response similarity, the *F*s being 35.13 and 17.21, respectively. An *F* of 6.84 is needed for significance at the .01 level. The interaction term did not approach significance. (Unless otherwise noted, all analyses of variance were based on distributions for which the variance was not significantly heterogeneous.) In terms of *t*s, using the within-groups error term (1.92), any difference in mean trials of 5.36 or larger is significant at the .05 level. While differences in learning as a function of stimulus similarity have frequently been reported, so far as is known this is the first clear demonstration of differences in learning as a function of response similarity. A previous study, while showing large differences

TABLE 1

Lists Used

(H refers to high similarity; L to low similarity; S to stimulus and R to response)

HS-HR	HS-LR	LS-HR	LS-LR
WEZ-Gleeful	WEZ-Agog	CEF-Gleeful	CEF-Agog
GYK-Genial	GYK-Neuter	MIB-Genial	MIB-Neuter
GOK-Pleasant	GOK-Pretty	GUK-Pleasant	GUK-Pretty
KIG-Smiling	KIG-Worldly	TIV-Smiling	TIV-Worldly
GUW-Blissful	GUW-Irksome	NUX-Blissful	NUX-Irksome
ZOW-Elated	ZOW-Vulgar	JUH-Elated	JUH-Vulgar
ZYW-Laughing	ZYW-Equal	ZER-Laughing	ZER-Equal

in error frequency as a function of response similarity, failed to show appreciable differences in trials to learn (**3**).

Error frequency reflects similarity differences almost perfectly; that is, the higher the similarity the greater the frequency of errors per trial. Although heterogeneity of variance prevents exact statements of significance, the *F*s for both response similarity and stimulus similarity are beyond the .01 significance level with interaction again being far from significant. In terms of *t*s, both comparisons (one with stimulus similarity low and one with it high) for influence of response similarity are significant beyond the .05 level. Comparable comparisons for stimulus similarity shows one *t* significant beyond the .05 level and one not significant. The latter occurs when response similarity is high.

Recall.—It will be remembered that recall of the stimulus items took place in two ways, first a controlled *drum* recall, and then an *unlimited* recall. The basic results are given in Table 3. For both drum and unlimited recall, two response measures are presented, namely, total correct syllables recalled and total number of letters correct, where correctness is defined as the correct letter in the correct position. As can be seen in Table 3, unlimited recall averaged 10–15% higher than drum recall; however, the correlations between these two types of recall ranged from .54 to .70.

TABLE 2

MEAN TRIALS TO LEARN AND MEAN ERRORS PER TRIAL

List	Trials		Errors	
	Mean	σ_M	Mean	σ_M
HS-HR	34.20	2.04	1.28	.12
HS-LR	28.53	1.65	.91	.07
LS-HR	25.17	2.21	1.02	.11
LS-LR	15.13	1.60	.66	.09

For drum recall, analysis of variance for syllables recalled shows only stimulus similarity to be significant, *F* being 6.73 with 6.9 necessary for the .01 level. The same variable approaches but does not reach significance for total letters recalled. The *F*s for response similarity and interaction between response and stimulus similarity fall far short of significance.

While Table 3 shows that for unlimited recall the mean differences are in the same direction as for drum recall, none of the variables achieves the .05 level. It is concluded, therefore, that stimulus similarity has a small but consistent effect on recall,

TABLE 3

MEAN NUMBER OF SYLLABLES RECALLED AND MEAN NUMBER OF LETTERS RECALLED ON DRUM AND WITH UNLIMITED TIME

List	Syllables				Letters			
	Drum		Unlimited		Drum		Unlimited	
	Mean	σ_M	Mean	σ_M	Mean	σ_M	Mean	σ_M
HS-HR	2.90	.35	4.07	.32	11.80	.92	15.80	.62
HS-LR	3.33	.30	4.40	.28	14.33	.84	16.33	.61
LS-HR	3.70	.32	4.40	.29	13.10	.74	16.20	.54
LS-LR	4.13	.27	5.00	.29	13.67	.74	17.53	.59

with the higher the similarity the poorer the recall.

It may be noted that in the case of total syllables recalled (drum) the over-all recall represents 50% of the items, while for letters the value is 61%. These values are considerably below those found by Thornton (83%) when he used adjectives for both stimuli and responses. Nevertheless, the present results essentially confirm the phenomenon that *S* does learn a number of R-S associations when asked to learn S-R associations.

Further analyses.—It has been shown that R-S learning does occur in paired-associate learning, but that it is not highly affected by stimulus or response similarity. It is assumed that if (as a control) *S* had never learned the S-R associations before being asked to recall the stimuli, he would have been unable to give any appreciable number correctly; certainly he would not have gotten 50% correct. On the other hand, if S-R recall had been taken following learning, it would probably have been 90–100% correct. The data thus indicate that R-S learning took place, but at a lower level than S-R learning. The relationship between R-S learning and other variables is now examined.

First, is R-S recall related to speed of learning the S-R associations? The *S*s in each group were divided into two subgroups based on speed of learning. The division was made on the basis of median trials to learn, so that 15 *S*s were in each subgroup. The fast learners were then compared with slow learners on R-S recall; none of the four *t*s approached statistical significance. Even allowing for differences in slope of S-R acquisition curves for fast and slow *S*s, the recall differences still would be small.

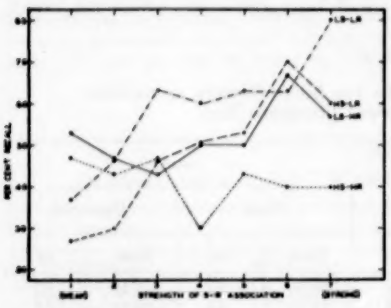

FIG. 1. R-S recall as a function of S-R strength. Strength of S-R association increases from Rank 1 through 7.

Recall of the S-R association has been shown to be directly related to number of correct anticipations during learning (4). The next question is, therefore, whether or not R-S recall is related to strength of S-R associations when the strength is measured by number of correct anticipations. The results of this analysis show that whether plotted separately for each list or whether all lists are combined, there was very little relationship. The recall of R-S associations started at about 50% when the items had been correctly anticipated twice (the lowest strength resulting from the criterion of learning used) and remained at about 50% irrespective of number of correct anticipations. This, then, is not in line with what has been found with S-R recall. However, this method of plotting recall as a function of strength is not very sensitive when wide differences are involved in trials to learn. Furthermore, all *S*s are not represented at all points. Another method utilizes the idea of the Vincent curve. For each *S* the strengths of the seven items for learning were rank-ordered, assigning Rank 1 to the item which had been given correctly the fewest number of times (weakest), and 7 to the item which had been anticipated correctly the greatest number of times (strongest). For each item it was noted whether or not the R-S recall was correct. Thus, the percentage correct at R-S recall can be plotted as a function of S-R strength and each *S* is represented once at each of the seven strengths. The results of these analyses are shown in Fig. 1. As can be seen, there is some relationship between recall of R-S and strength of S-R association for the two lists with

low response similarity. However, there is little if any relationship when response similarity is high. Looking at the data as a whole, R-S recall appears to be much less clearly related to strength of S-R learning than is S-R recall. At the present time there is no wish to speculate on these differences in R-S recall as a function of response similarity.

The fact that R-S recall is not strongly and unambiguously related to S-R strength will obviously require further work for clarification. If this is a form of incidental learning, it may be a simple and direct function of number of exposures of the S-R pairs. Since *Ss* were taken to a performance criterion in the present experiment, it cannot be determined if frequency of exposure (as contrasted to number of correct S-R anticipations) is related to R-S recall.

In view of Thornton's findings (discussed earlier) and the present results, the phenomenon of R-S learning seems to be a reliable one. The scope of the theoretical problem it represents can best be assessed after more is known of variables which affect it. It may well fit into the literature on incidental learning, and the generalizations which Postman and his students (e.g., **2**) are obtaining about incidental learning may also suffice for the phenomenon of R-S learning.

Summary

Subjects learned a list of seven paired associates. Immediately after learning they were presented the response items, one at a time, and were instructed to give the stimulus items associated with them during learning. The basic question was whether or not *S* could give the stimuli correctly. Does *S* learn an R-S association when instructed only to learn the S-R association? Four groups of 30 *Ss* each were used. Each group learned a different list. The four lists made up a 2×2 design in which stimulus similarity and response similarity were both high and low. For each list the stimulus items were nonsense syllables and the response items adjectives.

The results show:

1. In terms of syllables correct, recall of the stimuli (R-S) was 50% correct. In terms of correct letters in correct positions the recall was 61%.
2. While both stimulus and response similarity markedly effected rate of learning the S-R associations, only stimulus similarity produced significant differences on recall.
3. R-S recall was not strongly related to strength of S-R associations; when response similarity was high, there was virtually no relationship.

The results establish the existence of a phenomenon not hitherto reported but probably long suspected to exist. Whether or not it is a form of incidental learning requires further analyses.

References

1. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–269.
2. Postman, L., & Phillips, L. W. Studies in incidental learning: I. The effects of crowding and isolation. *J. exp. Psychol.*, 1954, **48**, 48–56.
3. Underwood, B. J. Studies of distributed practice: IX. Learning and retention of paired adjectives as a function of intralist similarity. *J. exp. Psychol.*, 1953, **45**, 143–149.
4. Underwood, B. J. Speed of learning and amount retained: a consideration of methodology. *Psychol. Bull.*, 1954, **51**, 276–282.
5. Underwood, B. J., & Goad, D. Studies of distributed practice: I. The influence of intralist similarity in serial learning. *J. exp. Psychol.*, 1951, **42**, 125–134.

(Received November 10, 1955)

Journal of Experimental Psychology
Vol. 53, No. 1, 1957

THE ROLE OF MEANINGFULNESS (*m*) IN PAIRED-ASSOCIATE VERBAL LEARNING[1]

CLYDE E. NOBLE AND DELDON A. McNEELY

Louisiana State University

A series of investigations (**4, 5, 6, 7, 8**) has been concerned with identifying and quantifying some of the variables which determine performance in verbal learning. In the first of these studies, Noble (**4**) developed a rational scale of meaningfulness (*m*) consisting of 96 dissyllabic nouns and paralogs ordered in terms of frequency of continued written associations in 60 sec. The *m* scale meets the need for an extensive range of verbal stimuli, from paralogs to frequently used words. It fulfills the requirement of high stability as well, having an estimated intergroup reliability coefficient of .993 ($n = 119$).[2] Recently, Mandler (**3**) has applied this technique to 100 English syllables ($n = 34$), while Umemoto, Morikawa, and Ibuki (**10**) have done the same for 1,892 Japanese syllables ($n \approx 100$).[3]

In his second investigation, Noble (**5**) found that variations in *m* produced significant effects upon serial verbal performance. Difficulty, measured by trials to mastery, was a decreasing curvilinear function of mean list *m* value, and there was an interaction between reactivity to *m* and *Ss'* initial ability. The relationship appeared to be negatively accelerated, but an exact test could not be performed with the *serial* task since only three satisfactory 12-item lists may be constructed from the sample provided by the *m* scale.

The purpose of the present experiment is to extend the analysis of meaning to the learning of *paired associates*. By controlling individual differences and eliminating the complicating effects of the ordinal positions of items, this method can evaluate more clearly than the serial method the nature of the difficulty-meaning relationship for specific S-R connections (**5**). The paired-associate task also permits the sampling of more empirical points than were available in earlier experimentation.[4] Specifically, it is predicted (*a*) that the acquisition rate of single verbal habits will be a positive function of *m*, and (*b*) that an interaction will occur between meaningfulness and ability to learn.

METHOD

Apparatus.—The apparatus consisted of a Gerbrands memory drum set to expose individual items at a 3-sec. rate. An idler attachment permitted five sets of 10 pairs to be presented.

[1] This paper is based in part upon a 1955 master's thesis investigation by the junior author under the guidance of the senior author. The latter reviewed the findings during a Symposium on "The Experimental Analysis of Verbal Behavior," Midwestern Psychological Association, May 1956. We are grateful to Drs. E. A. Bilodeau, G. A. Kimble, G. Mandler, and R. Thompson for helpful criticisms.

[2] Since the reported (**4**) average unit reliability of .975 was based on the *Z*-transformed intercorrelations of four groups of 30 *Ss* each, the combined scale reliability cited above represents a necessary correction for quadrupled sample size. The logic of this calculation as applied to response-defined variables is presented elsewhere (**9**).

[3] For kindly calling our attention to this article, we are indebted to Dr. Koji Sato of Kyoto University. The translation was performed by Mr. K. S. Kim.

[4] Two important studies on this topic (**1, 10**), which appeared after our work was completed, will be discussed later.

First, a practice list was constructed in which 10 pairs of three-syllable adjectives were arranged in five different random sequences. Each sequence constituted a trial, and the series repeated on Trial 6. The experimental lists were constructed from items in Noble's (4) *m* scale. Eighteen bands were prepared, each band having a list of 10 pairs of two-syllable nouns or paralogs. Like the practice list, these were also presented in five different random trial sequences giving effectively 90 sequences. The available tables of random numbers did not meet certain logical specifications, so the following statistical restrictions were imposed: no run was repeated on adjacent trials; no run was repeated more than once in five trials; there were no triple runs; and no S-R pair occurred in the same position on adjacent trials. Frequencies of occurrence at each serial position were approximately equalized for the 10 *m* values, and there were no homonyms, synonyms, antonyms, assonants, or alliterates within a given list. The purpose of these restrictions was to minimize cues which would facilitate the learning of some pairs and not of others.

Individual items ranged in median *m* value from .25 to 9.13. Medians were employed rather than means to provide greater accuracy at the low end of the scale, as discussed in (6). The net *m* value of each S-R pair was computed by averaging the median *m* values of the two items composing it. Pooling all 18 experimental lists, the paired associates represented *m* values ranging from .29 to 8.54 in 10 approximately equal-interval steps. The over-all mean was 4.29. There were from 3 to 7 items for each pair, with a total of 48 different items used altogether. No item appeared in more than 6 lists in the same S or R position, and none occurred more than 12 times. These manipulations were designed to stabilize performance scores at each of the 10 points by averaging out any intrinsic differential effects of particular items. All lists were typed in capital letters on white paper tape.

In order to test for the stability of the experimental variable, a 5-point graphic rating schedule of association value (a') was administered individually to an independent group of 10 undergraduates. Noble's (4) entire list of words and paralogs was used, and *Ss*' instructions were to judge the meaningfulness of each item on the basis of estimated number of associations. By a technique described previously (5, 6), 96 median ratings were derived. This scale had a high test-retest reliability ($r = .985$), and the correlation between test (a') and criterion (m) was .871. The evidence thus indicated that the *m* scale, although standardized on a military population, could be used quite rigorously with the population available in this investigation.

Procedure.—The *Ss* were 62 men and 28 women students at Louisiana State University, enlisted on a voluntary basis and uninformed of the purpose of the experiment. Their ages ranged from 18 to 27 (mean = 21 yr.). All were naïve with respect to paired-associate learning. Each *S* first learned the practice list to a criterion of 8 out of 10 correct responses, which was considered adequate to control the learning-to-learn factor in the present situation (11, p. 306). This was given during the same session as the experimental list and served to classify *Ss* into three pre-experimental ability levels (5). These levels, defined by the total trials required to reach the 8/10 criterion, were as follows: Slow > 17 ($n = 28$), Average 12–17 ($n = 32$), and Fast < 12 ($n = 30$).

The interval between onsets of successive stimuli was 6 sec.: 3 sec. for S alone and 3 sec. for S-R together. The intertrial interval was 9 sec. A 2-min. rest period was given between the completion of the practice list and the presentation of the experimental list. Five *Ss* were then assigned without bias to each of the 18 lists in a counterbalanced order of trial sequences. All 90 *Ss* were required to practice for 20 trials. On the first presentation of the list (Trial 1), *S* merely read the paired words aloud and standard pronunciations were established. Attempted anticipations of the correct responses began on the second presentation (Trial 2). The correction procedure was used.

For the analyses of variance a Type I design (2) was employed, in which the interaction of list and meaningfulness effects was confounded with individual differences. In this type of mixed factorial design inter-*S* comparisons are usually less precise than intra-*S* comparisons, hence individual differences were controlled in the case of *m* value and precision was sacrificed only in the ability and list evaluations.

Results

Figure 1 shows the percentage of correct responses as a function of practice with meaningfulness (m) as the parameter. The 10 original curves have been pooled by adjacent pairs for greater clarity of presentation. Each of the five categories in Fig. 1 thus denotes the mean median *m* value of two groups of 18 paired associates ($n = 90$). The functions are quite regular with no inversions until Trial 11. Consistent with previous data (5) from serial

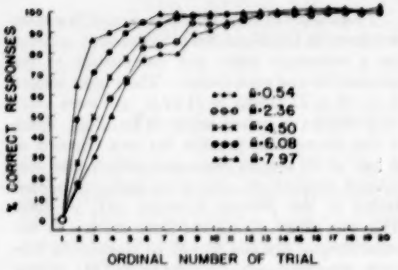

FIG. 1. Acquisition curves for single S-R pairs as a function of meaningfulness ($\bar{m}$).

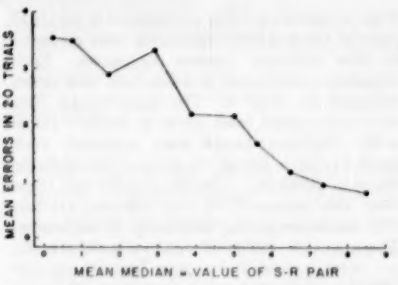

FIG. 2. The difficulty-meaning relationship for 10 groups of paired associates.

learning, rate of acquisition is a positive function of $\bar{m}$. Inspection of the forms of the individual curves reveals a trend toward positive acceleration with decreasing $\bar{m}$ value.

A 5 × 18 matrix of the average percentage of correct responses during Trials 2–20 was prepared. The A classification represented the same meaning categories as in Fig. 1, while the B classification represented the various lists. Each cell mean was based upon the performance of 5 *Ss*. The analysis of variance is presented in Table 1. Meaningfulness has a highly significant effect on performance, as indicated by the *F* ratio of 52.72 for $df = 4/288$. There are no significant differences among lists, which include individual differences. An interaction of marginal significance is present, but it is of secondary interest since the list classification was introduced primarily to formulate a more generalized estimate of the effects of *m* value. We may attribute the observed A × B interaction to experimental error associated with contextual distortion within certain lists.

TABLE 1

ANALYSIS OF VARIANCE OF PERCENTAGE OF CORRECT RESPONSES IN 20 TRIALS

Source	*df*	*MS*	*F*
Between *Ss*	89		
Lists (B)	17	311.88	1.20
Error (*b*)	72	260.34	
Within *Ss*	360		
Meaningfulness (A)	4	3046.72	52.72**
Interaction (A × B)	68	82.91	1.43*
Error (*w*)	288	57.79	
Total	449		

* $P = .05$.
** $P < .001$.

The difficulty-meaning relationship is shown in Fig. 2 where the mean total errors in 20 trials are plotted against the 10 average *m* values. A simple test of the observed trend is provided by computing a *t* ratio between the first and last points (**2**, p. 342). The ratio is 8.49, which for 89 *df* is highly significant ($P < .001$) of a tendency for errors to decrease with increasing meaningfulness. Another way of expressing the monotonic relationship in Fig. 2 is by means of rank-order correlation; the value of rho is − .988.

The correlation between the number of trials to 8/10 correct responses on the practice list and the same criterion measure on the experimental lists is .50, so we may expect a clear separation of the Slow, Average, and Fast groups for the foregoing analysis. Table 2 presents the total average errors in the various meaning categories for each ability level. It will be noted that the over-all means are

TABLE 2

MEAN NUMBER OF ERRORS IN 10 MEANING CATEGORIES FOR 3 PRE-EXPERIMENTAL ABILITY LEVELS DURING TRIALS 2–20

Levels	n	*m* value									
		.29	.79	1.74	2.98	3.92	5.08	5.65	6.51	7.40	8.54
Slow	28	4.64	4.71	4.57	4.79	2.89	2.86	3.00	1.68	1.46	.86
Average	32	4.28	3.38	2.53	3.56	2.66	2.09	1.38	1.31	1.06	1.03
Fast	30	1.73	2.47	1.67	1.67	1.07	1.57	.73	.57	.33	.47
Mean		3.55	3.52	2.92	3.34	2.21	2.17	1.70	1.19	.95	.79
σ_m		.312	.226	.298	.273	.187	.190	.202	.173	.167	.125

the same as the difficulty values graphed in Fig. 2. Standard errors of the means are included to show the inverse relationship between variability and meaningfulness. This rank-order correlation is − .915.

According to Table 1 there are no differences among lists, so the ability classification may be regarded as of primary interest. Pooling *S*s regardless of list, a 10 × 3 matrix of error scores was prepared. The analysis of variance given in Table 3 again indicates the facilitative effect of *m* on performance as well as differences due to initial ability. The significant interaction variance confirms Noble's (**5**) findings in the serial learning situation and supports his hypothesis that slow learners are more sensitive to differences in meaningfulness than are relatively fast learners. A Bartlett test indicated unequal variances among the three ability levels; for $df = 2$, $\chi^2 = 15.4$ which is significant ($P < .01$). Judging from an investigation by Norton (**2**, p. 78), however, such heterogeneity does not necessarily invalidate the *F* test, particularly since the main effects are still significant under a doubled level of confidence. The Bartlett test is useful here because it verifies the inverse relationship between ability and variability, demonstrating once again that sources of difficulty (e.g., low *m* value) affect the variances as well as the means in verbal learning.

In Fig. 3 appears an analysis of the difficulty-meaning relationship as a function of stage of practice. As was observed in Fig. 1, the acquisition

TABLE 3

ANALYSIS OF VARIANCE OF TOTAL ERRORS IN 20 TRIALS

Source	*df*	*MS*	*F*
Between *S*s	89		
Ability (B)	2	270.02	19.48**
Error (*b*)	87	13.86	
Within *S*s	810		
Meaningfulness (A)	9	100.59	25.66**
Interaction (A × B)	18	8.90	2.27*
Error (*w*)	783	3.92	
Total	899		

* $P < .01$.
** $P < .001$.

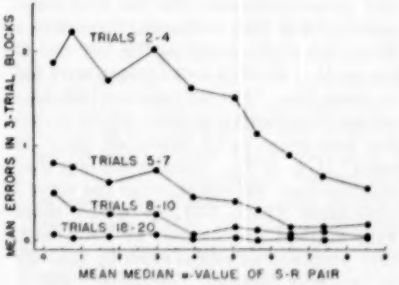

FIG. 3. The difficulty-meaning relationship as a function of stage of practice.

curves begin to level off around Trial 11 so the 3-trial parameter sequence is broken at that point. The errors for Trials 18–20 are plotted in order to describe the final learning period. From this graph it is evident that the phenomenon shown in Fig. 2 is primarily an initial effect which is quickly dissipated by the cumulative effects of practice.

Discussion

Two recently published experiments (1, 10) are closely related to the present findings. In conjunction with their application of the association-frequency method (4) to the Japanese language, Umemoto, Morikawa, and Ibuki (10) carried out a paired-associate investigation using three levels of meaningfulness ($n = 9$). Each list consisted of seven S-R pairs, with English letters as stimuli and scaled syllables as responses. Four different sequences were used and the items were presented at a 2-sec. rate. Consistent with the data of Noble (5) and with those reported here, difficulty varied inversely with *m* value: the mean trials to criterion for the low, medium, and high lists were 20.9, 13.9, and 7.5, respectively.

The other study, designed partly to test the same hypothesis, is by Kimble and Dufort (1). Employing 10-pair lists drawn from the *m* scale, they too obtained a generally negative relationship between trials to criterion scores and meaningfulness. In the first replication (Exp. IA), however, there was a flattening of the curve at the low end of the scale. Kimble and Dufort were led to state that "the paired-associate procedure introduces a variable which makes the less meaningful materials easy to learn" (1, p. 362). Despite the fact that this inversion did not occur in the second replication (Exp. IB), in which there were marked fluctuations at both ends, they proposed an explanation by analogy to the von Restorff "isolation" effect. They proceeded to ingenious demonstrations of other phenomena, such as heterogeneity in serial position curves (Exp. III), but unfortunately did not return to the central issue concerning us here. Kimble and Dufort seemed to be convinced that the difficulty-meaning relationship is nonmonotonic; i.e., it cannot be determined "in an uncomplicated way" (1, p. 361) by the paired-associate method.

The writers disagree on this point. Distortions resulting from a few different items appearing in diverse contexts can seriously obscure one's estimate of the true function. In executing the present experiment, therefore, a variety of S-R pairs were employed in a mixed factorial design. In addition, rather extensive statistical precautions regarding frequencies and runs were taken in order to preclude other kinds of artifact. There was also an attempt to reduce variability due to learning-to-learn, type of score, and sample size. Whereas Kimble and Dufort used 40 *S*s with no pretraining, 2 lists, 12 different sequences, and a trials-to-mastery design, this experiment used 90 *S*s with pretraining, 18 lists, 90 different sequences, and a constant-trials design. The two studies are very similar otherwise, even to the extent of identical apparatus and presentation rates. Under the present conditions there is no reason to conclude that paired associates per se introduces an "isolation" factor. While it is true that a list × meaning interaction significant at about the 5% level was obtained, the effect is confounded with individual differences; indeed, this may help to clarify the Kimble-Dufort observations by identifying the mode of operation of the complicating factor.

More to the point are the relative magnitudes of the meaning and interaction variances. Given the precision indicated by Tables 1 and 2, the practical importance of the "isolation" effect is small when compared with the main effect of meaningfulness (2, p. 214). Furthermore, since no significant differences were found among lists, the generalization of this interaction to any hypothetical population of *S*s would

have to be so limited by qualifications as to be of small theoretical value. Accordingly it is interpreted as experimental error. In other words, the phenomenon Kimble and Dufort observed may be an artifact and not representative of the function one would obtain by infinite replications where items, positions, and sequences are of no interest.

It is notable that response variability in these studies is rather large. This may be due as much to the *material* as to the *method*. Perhaps if more uniform stimuli (e.g., three-letter syllables vs. two-syllable words) were employed, thus requiring more uniform responding (e.g., spelling vs. pronouncing), the resulting functions would show greater stability and freedom from disturbing artifacts. Kimble and Dufort's own evidence (1, p. 363), collected during preliminary investigation of the comparative effects of S and R meaningfulness (Exp. II), is relevant here. Before this hypothesis can be tested, however, new measurements of association value will be required. Lists such as those compiled by Glaze and Krueger have been very useful to students of verbal learning, but are now outdated. Mandler's (3) scale, which is similar to the *m* scale operationally, might be suitable except that the range of syllable associations (2.9 to 5.3; $M = 4.1$, $SD = .48$) is more restricted than that of paralogs and words (.99 to 9.61; $M = 3.66$, $SD = 2.38$).

The present results suggest that the difficulty of single verbal habits is a decreasing monotonic (sigmoid?) function of *m*. The theoretical explanation for the facilitative effect, however, is not clear. There have been a number of attempts to account for the meaning-learning correlation; various writers have alluded to such factors as "familiarity" (Ebbinghaus), "multiple associations" (James), "organization" (Koffka), "predifferentiation" (Gibson), and "transfer" (McGeoch). Actually, these terms are little more than descriptions of some of the related phenomena. A more strategic attack upon the problem would be to determine the relative contributions of the S and R terms to the variance in performance, as outlined by Noble (5, p. 443). Several lines of evidence reviewed elsewhere (5, 8) indicate that differential facilitation is primarily connected with responding (articulation) and that stimulus factors are of secondary importance (cf. also 10). Consideration of those studies together with the current discussion permits two final predictions: (*a*) the simple effect favoring high S-high R meaningfulness (e.g., *m* = 6–9) over low S-low R meaningfulness (e.g., *m* = 0–2) will be greater than for any other comparison, and (*b*) the difference in the main effect of *m* on the response side will be greater than the comparable difference on the stimulus side. Given (*a*) and (*b*), the likelihood of intrinsic S-R interaction with respect to *m* is slight.

Both these hypotheses are clearly at variance with the Kimble-Dufort position, yet there is one point of agreement. By assuming that "isolation" is a perceptual process, they reason (1, p. 363) that the learning of low S-high R pairs should be enhanced relative to that of high S-low R pairs. It is interesting that an S-R motor-patterning viewpoint, although contradictory otherwise, generates the same implication. It is also consistent with the methodological consideration that *m* values are established by *recall* rather than *recognition* operations. Thus, using Fig. 2 as a guide, the expected ranking of the four basic conditions in order of decreasing difficulty is: 1. low S-low R, 2. high S-low R, 3. low S-high R, 4. high S-high R. By comparison, the configurational ground plan is at least incomplete for it neglects to establish continuity with the limiting cases, where the *m* values of S and R covary. Due to the absence of a low S-low R condition in the Kimble-Dufort design (Exp. II), and also because of the presence of confounded (and possibly heterogeneous) interaction, their data do not constitute an appropriate test of the hypotheses presented above.

A simple 2 × 2 factorial design employing four independent groups of *Ss* would provide unequivocal answers to these questions.

Summary

Certain relationships between performance in verbal learning and the attribute of meaningfulness (*m*) were investigated. Ninety college students practiced for 20 trials with 18 lists of 10 paired associates representing approximately equal intervals on the *m* scale. As expected from serial learning research, rate of acquisition was a positive function of *m*, and the curves showing the percentage of correct responses tended toward positive acceleration with decreasing *m* value. Difficulty measured by mean total errors was an inverse monotonic function of *m*, the rank-order correlation being − .988. Variability also decreased with increasing meaningfulness.

There were significant differences in performance correlated with pre-experimental ability, as well as a negative relationship between ability and variability. Consistent with a second prediction, reactivity to *m* interacted with ability to learn. A level of practice analysis revealed that the difficulty-meaning relationship was an initial effect which disappeared by Trial 11. Contrary to recent opinion, these findings indicate that the law relating difficulty and meaningfulness for specific S-R connections can be determined by the method of paired associates. Suggestions were offered for further research to specify more clearly the role of "perceptual" and "motor" factors in verbal learning.

References

1. Kimble, G. A., & Dufort, R. H. Meaningfulness and isolation as factors in verbal learning. *J. exp. Psychol.*, 1955, **50**, 361–368.
2. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* New York: Houghton Mifflin, 1953.
3. Mandler, G. Associative frequency and associative prepotency as measures of response to nonsense syllables. *Amer. J. Psychol.*, 1955, **68**, 662–665.
4. Noble, C. E. An analysis of meaning. *Psychol. Rev.*, 1952, **59**, 421–430.
5. Noble, C. E. The role of stimulus meaning (*m*) in serial verbal learning. *J. exp. Psychol.*, 1952, **43**, 437–446; 1952, **44**, 465.
6. Noble, C. E. The meaning-familiarity relationship. *Psychol. Rev.*, 1953, **60**, 89–98.
7. Noble, C. E. The familiarity-frequency relationship. *J. exp. Psychol.*, 1954, **47**, 13–16.
8. Noble, C. E. The effect of familiarization upon serial verbal learning. *J. exp. Psychol.*, 1955, **49**, 333–338.
9. Noble, C. E. Scale reliability and the Spearman-Brown equation. *Educ. psychol. Measmt*, 1955, **15**, 195–205.
10. Umemoto, T., Morikawa, Y., & Ibuki, M. The non-association values and meaningfulness of 1892 Japanese two-letter syllables and words. *Jap. J. Psychol.*, 1955, **26**, 148–155.
11. Underwood, B. J. *Experimental psychology.* New York: Appleton-Century-Crofts, 1949.

(Received June 3, 1955; Revised July 5, 1956)

Journal of Experimental Psychology
Vol. 53, No. 1, 1957

A RE-EXAMINATION OF RUBIN'S FIGURAL AFTEREFFECT

IRVIN ROCK

New School for Social Research

AND

IRWIN KREMEN[1]

Harvard University

There are very few experiments which unequivocally show form perception to be determined by specific past experience, if by form perception we mean the experience of shape as such. By this definition we mean to rule out influences of past experience which lead to recognition of a form (familiarity) and meaning, which obviously do occur. Leeper's experiment showing a role of past experience in the correct organization of Street Figures does seem to be a case of past experience influencing form perception (**4**). Another is the Schafer and Murphy experiment (**8**), although it now is not certain under what conditions their results can be reproduced [(**6**) but see also (**3**) and (**9**)[2]]. But perhaps the clearest example of a reported influence of this kind is Rubin's famous figural aftereffect (**7**). Rubin first showed *Ss* ambiguous figure-ground patterns, instructing them to see each in a certain way. Subsequently these figures were exposed again briefly with instructions to report what was seen. He found a beyond-chance tendency for *Ss* to organize the figures in the way they had seen them previously. This past-experience effect on figure-ground *perception* is not to be confused with the other fact reported by Rubin, namely that *recognition* of these figures depends upon whether or not they are seen in the same way when re-exposed. We are here concerned only with the first effect.

There is a logical reason, however, why past experience should not determine form perception as here defined. Insofar as past experience with specific shapes is preserved via memory traces, to say that such past experience can determine form is to say that the relevant trace can enter into the process which organizes the percept. But we assume that ordinarily the relevant trace is aroused *after* the form is perceived, that it is selected in some way by virtue of the similarity of the present perceptual process to the trace and that this leads to recognition (**10**). Before the form is perceived what reason is there for any particular trace to play a determining role, particularly in the light of the fact that the image of the form will rarely fall again in the same place on the retina? This reasoning applies to Rubin's experiment because here a past-experience effect means that the trace (or traces) of any given figure seen in training must make its

[1] The authors wish to express their appreciation to Professor Milton Blum and the 1955 Summer Session Faculty of the Psychology Department, 23rd Street Division, City College of New York, for their cooperation in placing facilities and *Ss* at their disposal. They also wish to thank Professor Mary Henle for her careful reading of the manuscript.

[2] Recently Prentice and Taylor repeated Smith and Hochberg's experiment with negative results (**5**).

FIG. 1. (A) One of Rubin's figures; the surrounded part was green. (B) The black training figure corresponding to Test Figure 13. (C) The white training figure corresponding to Test Figure 13.

influence felt *before* the ambiguous stimulus is seen one way or the other in the test, which means before one region or the other emerges as the shaped figure. Hence if such effects do occur one has the problem of explaining how they are possible.

In reading carefully Rubin's procedure, however, it was concluded that it contained many serious flaws so that one could not really have confidence that the figural aftereffect actually occurs. Gottschaldt repeated Rubin's experiment and reported that unless *S*s expect a test and are set to look for familiar figures during the test, the effect does not occur (**2**). Since in Rubin's procedure these attitudes were undoubtedly present it might be argued that without them he too would have obtained negative results. But in any case Gottschaldt followed Rubin's procedure in all other respects so that it was still considered advisable to repeat it in such a way as to remove the questionable aspects of the procedure and to give the effect optimum conditions under which to show itself.

Rubin's procedure will be outlined and the questionable features indicated. He used irregular forms of the kind shown in Fig. 1A, in which the surrounded figure was cut out of cardboard and measured about 1.5 cm. square. When placed before a lantern with a green glass plate interposed they were projected on a screen and could be seen as surrounded green shapes or surrounding black shapes. Two sizes of projected figures were used, either true size 1.5 cm. square or enlarged size 7.5 cm. square (.6 in. square and 3 in. square, respectively). The *S* sat at a distance of 60 cm. in front of the screen with the projection apparatus behind it.

Rubin worked with three *S*s over a period of seven days. Each day *S* was first presented with a *training series* consisting of two divisions of nine figures each, in one of the sizes employed. In the case of one division, *S* was instructed to try to see the enclosed portion of each card as figure; for the other, he was instructed to see the surrounding portion of each as figure. The nine figures of each division were shown for 4 sec. each, and the series was presented four times over. Following this, the entire procedure was repeated for the other size figures employed, thus making a total of 36 training figures shown per day, each for a total of 16 sec. From day to day each *S* was given new figures in the same way.

Following a 30- to 45-min. intermission a *test* was given each day. Rubin shuffled the 36 training figures with 18 new ones and presented them in random order. The *S*s were instructed to remain passive and to report whether the enclosed or the surrounding portion appeared as figure and also whether or not the figure seemed familiar (by adding these instructions, Rubin was able to obtain data concerning the recognition effect). In the test, *S* was given a ready signal and the lantern was uncovered briefly for each exposure; it was covered again as soon as *S* responded.

The following features of this procedure seem questionable from the point of view of testing the hypothesis under consideration. As Gottschaldt subsequently pointed out, Rubin's *S*s undoubtedly expected that the test figures would contain the training figures because they were asked to report not only which figures they saw but whether they appeared familiar. The latter instructions may even have led *S*s to believe that

they were *supposed* to see familiar figures. In addition, the use of so many figures per day was inadvisable because, even without such instructions, *S* might begin to recognize figures after a good number had been shown, so that for the remaining ones he might develop a similar expectation. Worse still is the use of *S* for seven days of experimentation; after the first day he is completely sophisticated concerning the entire procedure. In fact it is not clear from Rubin's description whether the *S*s were initially naïve concerning the hypothesis. Obviously, it would be preferable to use many *S*s, each for one cycle of training and test, instead of a few for many such cycles.

Another shortcoming of Rubin's procedure is that his figures were not truly ambiguous, the enclosed portions being slightly favored. His results for the 18 new figures shown during the test each day reveal a dominance of enclosed portion over surrounding portion of about 60:40. However in a repetition of this experiment dealing only with the recognition effect, Dutton and Traill (1) found that most *S*s experienced great difficulty in seeing the surrounding portion as figure. It would therefore have been of interest to have control data from *S*s without any previous training with the critical test figures.

Still another major weakness of Rubin's procedure stems from the manner in which *S* obtained experience with the training figures. The *S* was told to see the figures as instructed but he did not always succeed in doing so. Rubin, therefore, distinguishes "obedient" cases where *S* did succeed from "disobedient" ones where he did not. This difficulty may actually be unfavorable from the standpoint of obtaining the effect because not all of the time allotted to each training figure is necessarily spent perceiving the intended organization. In any case it is unsatisfactory not to be sure what *S* will see in training.

Finally, Rubin apparently did not control exposure time during the critical test. One would judge that the time interval was not too short because it ceased only when *S* responded. This is a serious objection because if this is the case *S* may have had time to reverse the figure.

Method

In general, changes were introduced either because it was thought they would facilitate the aftereffect or because they were deemed necessary to make sure that the effect, if obtained, could be attributed to the spontaneous influence of past experience on figure-ground organization.

After exploratory experimentation, 18 ambiguous black-white figures plus one additional sample were prepared (see Fig. 2). They were 2 in. square, mounted on neutral gray cardboards (TV illustration board No. 350), 5¾ in. × 6¾ in. A thin border line of black extended around the white fields. This center-contour type of figure-ground pattern was selected because it is more genuinely reversible. The white field on one side could easily be seen as figure against the gray ground and the black field on the other side could also easily be seen as figure against the gray ground. Of the 18

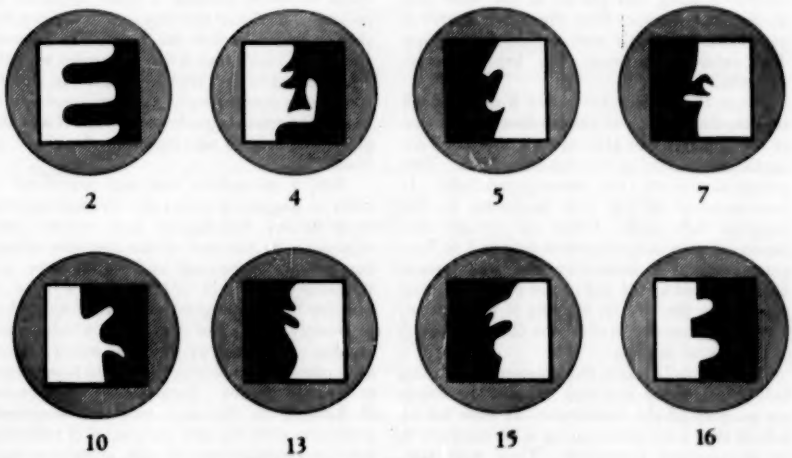

FIG. 2. The eight critical test figures.

figures, 9 had black fields on the left and 9 had black fields on the right. All were nonsense figures.

Eight figures, those which yielded a fairly even division of black and white responses when shown to a control group, were selected for use as critical test figures. From these, eight training figures were prepared for Group A consisting of only one-half of the original figures (Set A). They were thus unambiguous white or black figures on the same gray ground as were the test figures (see Fig. 1 B and C). Four were black and four were white, with two of the black having the contour on the left, two having it on the right, and similarly for white. The complementary halves of the test figures were prepared as training figures for Group B (Set B). This type of training figure was used so that there would be no doubt that *S* experienced just the half of the critical figure intended during training. Only eight critical figures were used, separated in the test by completely new ones, in order to prevent *S* from becoming aware, soon enough for it to influence the results, that previously seen training figures were being shown in a new setting.

The apparatus consisted of a modified Dodge Tachistoscope, the arms of which measured 38 in. With *S*'s head resting against the apparatus so that he was looking down one of the arms and the figure placed in a slot at the other end, *S* was 38 in. from the figure. A variable shutter was placed inside one of the arms with a plunger extending outside. It was connected to the light sources in each of the two arms. A cardboard with a circular opening equal in size to the shutter opening was placed in the other arm at the same distance from the corner mirror so as to coincide exactly with the shutter opening. The experimental room was kept in semi-darkness.

When the shutter was closed *S* saw through the circular opening of the cardboard a portion of the homogeneous gray surface of a 5 × 7 in. cardboard inserted at the end of one arm. This served as the pre- and postexposure field. It was matched in hue and brightness to the exposure field itself. When the plunger was depressed, part of the exposure field (a 5 × 7-in. gray cardboard) containing the 2-in. square figure inserted at the end of the other arm, was seen within the circular opening of the shutter. The *S* used one eye to eliminate double imagery of the shutter opening.

Training.—To allow the experience with the half-figures to be acquired without disclosing the purpose of the experiment, *S*s were led to believe they were participating in an experiment on extrasensory perception. They were told to look at each figure for as long as it was exposed and attempt to "project its image to another person in another room." The eight training figures were then each exposed for 2 sec., eight times over, the order being randomized for each complete presentation.

Intermission.—In order to prevent any expectation of a test from developing, *S*s were then asked to construct a design from little colored pegs for the next 5 min.

Test.—The eight critical test figures were presented together with 10 new figures of the same type (5 black on right, 5 black on left). The order of the 18 cards was arranged in such a way that any two critical cards were always separated by at least one new one. The eight critical figures were so arranged that a figure for which the training figure had been white was followed by one for which it had been black, etc. The color arrangement of each figure (i.e., whether black or white is on the left side) was alternated from figure to figure from the first to the eighteenth.

Each card was exposed for 1 sec., this exposure time being selected after an initial failure to obtain the effect with shorter exposures. It was decided to allow 1 sec. inasmuch as Rubin did not control his time interval and it may have been that long. A ready signal was given prior to each exposure.

The *S* was first given a brief explanation of figure-ground perception by means of a sample figure. He was made to understand that each test figure could be seen in either of two ways and that *E* was interested in knowing whether in any given card the black or the white portion appeared to him as the figure. No request was made to report whether a figure appeared familiar. The *S* was also required to indicate the side on which he saw the figure. This served as a rough check that *S* had responded with the color he had in fact seen as figure. In this respect there were virtually no errors among all *S*s in all the variations performed. The *S* was also given the option of reporting "neither" or "both."

Also, a recognition test was introduced in order to demonstrate that the training with the unambiguous half-figures had indeed been effective. At the end of the test the critical figures were presented again (but now not tachistoscopically, in order to allow time to examine the two alternatives) and *S* was asked to identify the half of each figure which looked familiar. Evidence of retention would require more correct recognitions than could be expected to occur by chance. To prove the effectiveness of learning in this way may be considered gratuitous since the only purpose of the training was the establishment of past experience with the half-figures. It may be assumed that these

were perceived, leaving traces of themselves, and that this is sufficient for an influence on later perception even if subsequently they cannot be consciously identified. Nevertheless, the objection might be raised, and this would apply to Rubin's experiment as well, that the training had not been sufficiently extensive to expect such a past-experience effect to occur. This objection is met if the training was sufficient to enable *Ss* to identify the training figures even after the critical test.

Control group.—These *Ss* were given 1-sec. exposures of the 18 figures in the test series without any prior training and in exactly the same way as the experimental *Ss*. This provided data concerning the way the critical figures tended to be organized without prior experience with either half of the figure-ground patterns.

Subjects.—An approximately equal number of male and female undergraduate students were used in all groups. There were 15 *Ss* in Group A, 15 in Group B, and 20 in the control group.

To summarize, the present experimental procedure differed from Rubin's in the following ways: (*a*) The figures differed in the direction of being more genuinely reversible. They were the center-contour type of ambiguous figure. (*b*) The training figures were half-figures which insured that the experience was with the half intended. Rubin used ambiguous figures with instructions to see each in a particular way. (*c*) The figures were 2 in. square at a distance of 38 in. from *S*. Rubin's figures were either .6 in. square or 3 in. square at a distance of 60 cm. (24 in.) from *S*. (*d*) The projection apparatus differed primarily in that it provided a pre- and postexposure field while Rubin's did not. (*e*) In several ways the procedure eliminated expectation of a test which might lead *Ss* to look for familiar figures: the instructions deleted the request to report whether a figure appeared familiar; naïve *Ss* were used for only one cycle each rather than a few *Ss* used for a complete cycle of the experiment on seven successive days; the training session was disguised as an extrasensory perception experiment. Considerably fewer figures were used than in Rubin's experiment. (*f*) Exposure time was limited to 1 sec. by tachistoscopic presentation. (*g*) Training consisted of eight 2-sec. exposures of each figure rather than of four 4-sec. exposures. (*h*) The intermission period was shortened to 5 min. (*i*) A control group was included. (*j*) A recognition test was introduced following the main test.

Results

Of the 18 figures shown to the control group, the black and white halves were seen about equally frequently in eight. Consequently these were selected as critical figures. Four had black on the left and four had black on the right. For these eight critical figures there were altogether 76 responses to Set A halves and 71 responses to Set B halves (Table 1).

TABLE 1

Responses to the Critical Figures

Critical Figure	Group					
	Control (N = 20)		A (N = 15)		B (N = 15)	
	Set A	Set B	Training Half (Set A)	Non-Training Half (Set B)	Training Half (Set B)	Non-Training Half (Set A)
2	11 (B)	9 (W)	12	2	6	7
4	10 (W)	10 (B)	5	8	11	4
5	10 (B)	8 (W)	7	5	10	4
7	10 (W)	10 (B)	6	8	7	8
10	8 (B)	9 (W)	7	7	6	8
13	8 (W)	8 (B)	9	6	6	6
15	9 (B)	9 (W)	4	9	10	5
16	10 (W)	8 (B)	6	6	7	6
Sum*	76	71	56	51	63	48

* "Neither" and "both" responses are not included.

The mean number of correct identifications obtained in the recognition test for the 30 *Ss* of Groups A and B was 6.0 with an *SD* of 1.3 (the mean for incorrect identifications was 1.5 since *S* did not make a choice in every case). This value differs significantly from chance at the .001 level. The experience with the training figures was sufficient to allow for subsequent recognition. It seems, therefore, that this training should be sufficient to allow for a past-experience effect on figure-ground organization if there is such an effect.

As shown in Table 1, in the critical test Group A gave 56 responses to Set A halves and 51 responses to Set B halves, and Group B gave 63 responses to Set B halves and 48 responses to Set A halves. As one can see by inspection, the differences between the distribution of results for Group A and the 76:71 distribution for the control group is not significant. Nor is the difference between the distribution of results for Group B and the 71:76 distribution for the control group significant ($\chi^2 = 1.6$, $P < .3$ for 1 *df*). Comparison of the total responses to those figures seen in training (119) and those not seen in training (99) for both experimental groups combined, also reveals no significant effect ($\chi^2 = 1.8$, $P < .2$ for 1 *df*).

The following variants of the analysis of these data were used, always with the conclusion that no significant effect resulted from the training: (*a*) consideration of the results only for *Ss* having very high recognition scores; (*b*) elimination of the results of the few *Ss* who responded with what appeared to be either positional sets or color sets; (*c*) comparison of the results in terms of the number of *Ss* seeing more training than nontraining against the number seeing more nontraining than training halves; (*d*) determination of whether the number of *Ss* who do seem to show the effect are in excess of chance.

The results reported above duplicate findings obtained in an earlier version of this experiment, the procedure of which differed in the following respects. A projection-type of tachistoscope was used which cast an image 9½ in. square on a screen 84 in. from *S*. The critical test figures were not selected on the basis of the control data. The eight training figures were each exposed for 4 sec., four times over. The intermission lasted 10 min. during which time *S* identified Street Figures. A recognition test was not included. The results of this experiment with 30 *Ss* also did not reveal an effect of past experience.

Discussion

Although negative results were obtained, it is impossible to establish conclusively that an effect such as this cannot be obtained under *any* circumstances. On the other hand, it seems correct to say that the failure to duplicate Rubin's results shows that the type of past-experience effect he had in mind does not occur with a procedure designed to reveal it. If the effect requires very special conditions to show itself then this necessitates a drastic change in the evaluation of its significance.

It is not known whether an effect of past experience would have been obtained if the present procedure were modified by giving *S* the additional instruction to report whether a figure appeared familiar. But an effect which requires a set such as would no doubt be induced by these instructions could not be regarded as evidence of a spontaneous influence of past experience on figure-ground perception. Under the influence of such a set *Ss* might tend to look for the familiar half. Even in a fairly brief exposure there would be time to organize first one and then the other half as figure. The one recognized would probably then be reported. This would be evidence of an effect of past experience on recognition but not on perception.

It may be objected that the use of the unambiguous halves of the ambiguous figures in training, rather than the ambiguous figures themselves under instructions to see them in a particular way, changes the *kind* of past experience *S* gets. In Rubin's procedure *S* sees an *ambiguous* training figure in a particular way and the question is, will he tend to see it again in the same way? In the present procedure, *S* sees an *unambiguous* training figure and the question is, will he tend to see this figure when an ambiguous pattern, containing it as one possibility, is shown? A first answer to this objection is that there seems to be no practical way of using an ambiguous figure in training while guaranteeing that *S* will see it in the desired way, and only in this way, each time it is shown. A second answer is that the organized percept is presumably identical whether one sees an unambiguous figure or whether one sees as figure that same geometrical shape in an ambiguous figure-ground pattern.

In a repetition of the Schafer and Murphy experiment by Rock and Fleck it was pointed out that for many *S*s the training faces appeared somewhat different when seen in the ambiguous test pattern (**6**, p. 774). In the present experiment, however, there was no evidence of this, and there is the excellent recognition test results as proof that experience with the half-figure sufficed for recognition of it when the full figure was exposed.

The results of this experiment can also be viewed in relation to the problem of the influence of motivation on perception. It is likely that if motivation influences perception it does so through the influence of traces representing the need-related objects (**10**). Consider the Schafer and Murphy experiment which structurally is very much like Rubin's experiment. In the former, there are also two alternatives in the ambiguous figure-ground test situation. The perception of one has been previously rewarded and the other punished. If, however, reward and punishment is to have an effect, it can only be via the traces of these training figures. One might say, for example, that the trace of the previously rewarded figure is somehow stronger or more strongly cathected or dominant in some way in comparison with the trace of the previously punished figure. Thus the rewarded traces can be expected to be more effective than the punished ones. But in the Rubin experiment the demands made for the appearance of the effect are simpler. Here there is a trace for one alternative and *none* for the other. If the effect does not occur in the trace vs. no-trace situation why should it be expected to occur in the cathected vs. noncathected trace situation? Hence if this effect is confirmed in subsequent investigations (in view of the contradictory results with differing conditions thus far) we will be faced by a puzzling problem of interpretation.

Summary

Rubin's experiment showing a tendency to organize ambiguous figure-ground patterns in accordance with past experience was repeated with certain modifications. Questionable features of the original procedure were eliminated. The effect was not obtained.

References

1. Dutton, M. B., & Traill, P. M. A repetition of Rubin's figure-ground experiment. *Brit. J. Psychol.*, 1933, **23**, 389–400.
2. Gottschaldt, K. Über den Einfluss der Erfahrung auf die Wahrnehmung von Figuren, I. *Psychol. Forsch.*, 1926, **8**, 261–317.
3. Jackson, D. N. A further examination of the role of autism in a visual figure-ground relationship. *J. Psychol.*, 1954, **38**, 339–357.
4. Leeper, R. A study of a neglected portion of the field—the development of sensory organization. *J. genet. Psychol.*, 1935, **46**, 41–75.
5. Prentice, W., & Taylor, M. Personal communication, July, 1955.

6. Rock, I., & Fleck, F. S. A re-examination of the effect of monetary reward and punishment on figure-ground perception. *J. exp. Psychol.*, 1950, **40**, 766–776.
7. Rubin, E. Visuell wahrgenommene Figuren. Copenhagen: Gyldendalske, 1921.
8. Schafer, R., & Murphy, G. The role of autism in a visual figure-ground relationship. *J. exp. Psychol.*, 1943, **32**, 335–343.
9. Smith, D. E. P., & Hochberg, J. E. The effect of "punishment" (electric shock) on figure-ground perception. *J. Psychol.*, 1954, **38**, 83–87.
10. Wallach, H. Some considerations concerning the relation between perception and cognition. *J. Pers.*, 1949, **18**, 6–13.

(Received December 20, 1955)

Journal of Experimental Psychology
Vol. 53, No. 1, 1957

EXPERIMENTS ON FIGURAL DOMINANCE[1]

THEODOR M. KÜNNAPAS

University of Stockholm, Sweden

Rubin (**8, 9**) has pointed out that a small enclosed field will much more easily be perceived as figure than a larger surrounding field. Harrower (**3**) found that smallness of area and surroundedness, closure, good continuation, articulation, and asymmetry are the most important factors which determine the figure-ground articulation. Graham (**2**) was the first to attempt quantitative measurement of the influences of the area and the color in a reversible configuration. He found that the smaller fields are more dominant, and this dominance varies inversely with area. Observations of Köhler (**6**) agreed with those of Graham. Using a modified form of Dodge tachistoscope, Goldhamer (**1**) substantiated Graham's conclusions. Oyama and Torii (**7**) analyzed the effects of area, voluntary control, and prolonged observation by the continuous presentation method trying to obtain several quantitative measures.

In the present paper an attempt is made to study the relative dominance of the alternating figures in a reversible pattern. It may be called figural dominance. The area of the figure cannot be the primary factor in figural dominance. Behind the area is one or another principle of organization. In the case of Köhler's circle with six sectors, which was used by Graham and others, the principle of proximity is probably an important organizing factor. As Koffka (**5**, p. 166) says, "two parts in the field will attract each other according to their degree of proximity." The nearer pairs spontaneously unite, and the shorter the intrafigural distance, as compared with the interfigural, the more stable and dominant will the unit be.

This study represents an attempt to determine quantitatively the extent to which the principle of proximity in a reversible figure-ground pattern will influence the figural dominance. Instead of Köhler's figure, Rubin's pattern with eight sectors was employed since it is more symmetrical and more easily fixated by *O*s.

METHOD

Stimulus material.—The figures used as stimuli were similar to the drawing in Fig. 1. In each figure the circle (diameter = 8 cm.) was divided into eight sectors by black lines drawn with India ink. Four of these sectors lying in vertical and horizontal positions formed the "plus" figure, the other four in the tilted positions formed the "cross" figure. These alternative figures had no differences in brightness nor in color. The whole pattern was drawn in the center of a white card which was 25 cm. square. Nine figures were used. The angles of the plus-figure varied as follows: 5°, 15°, 25°, 35°, 45°, 55°, 65°, 75°, and 85°; the angles of the cross-figure were, accordingly, 85°, 75°, 65°, 55°, 45°, 35°, 25°, 15°, and 5°.

Apparatus.—Alternating figures of the stimulus pattern were recorded by means of L. M. Ericsson's Morse-writer with a telegraph key. The key was operated by *E*, who marked the moment of the beginning of each reversal on a strip of prepared paper which had a constant speed of 34 mm. per sec.

Procedure.—In front of *O* was a support for the stimulus card with a variable inclination. By means of this support, the card was placed orthogonal to *O*'s line of vision. The distance between *O*'s eyes and the pattern was 50 cm. corresponding to a visual angle of 9°.

The *O* was instructed as follows: "You are to see this reversible pattern. Shut your eyes. Two seconds before the trial, *E* will say 'Ready,' and at the word 'Now' open your eyes and fixate

[1] The author is greatly indebted to Professor Gösta Ekman for his valuable advice.

FIG. 1. One of the stimulus patterns used in the experiments. Here, the plus-figure has 35° and cross-figure 55°.

the center of the pattern where the lines meet, and keep the fixation throughout the entire trial. Try to reduce during the fixation winks of the eyes as much as possible. Say 'plus' as soon as you see the plus-figure as figure and 'cross' when the cross-figure is figure. Begin with the first figure you see. Say 'No' whenever you cannot see any figure or you have a somewhat diffuse, unclear, vague perception or you could not report the reversal at the right moment. Take a passive attitude and do not try to influence shifting."

Observations were made with both eyes. In preliminary trials *Os* were made familiar with the experimental situation. In order to make observation as simple and natural as possible to *O*, the reversals were recorded by *E*, as described above. The *E* also recorded, for each trial, whether plus or cross was seen first as "figure."

Four recording series were taken for each of the nine figures on two different days: two series were given each day. A series consisted of two trials, or more when it was difficult to observe the shiftings. Each trial lasted at most 30 sec. Every time *O* reported that he did not observe clearly a figure, a short pause was immediately introduced. During the pauses of 1 min. between two trials *O* kept his eyes shut.

During one series at least 15 clearly seen figural periods were recorded. Thus, for each *O* there were 60 figural periods recorded for each of the nine patterns. In the 5° condition for the plus-figure, however, the dominance of the plus-figure was already so great that for some *Os* it was impossible to record altogether 60 figural periods. Sometimes they did not see reversals at all. Therefore, for three *Os* who apparently reached the limit of reversals there were recorded less than 60 figural periods ($O_4 = 31$, $O_5 = 14$, $O_{10} = 14$ periods).

There were also for each *O* and each pattern eight records in which the plus or cross was seen first as figure, at the beginning of each of two first observation trials during a series.

The sequence of the patterns was rotated so as to equate any possible learning and fatigue effects.

Observers.—Eleven students of psychology, six men and five women, took part in the experiments. All *Os* had acquired a certain practice in the observation of the reversals during the laboratory course in experimental psychology. All *Os* had normal acuity, in one case corrected by glasses.

RESULTS

There are at least five measures related to aspects of figural dominance: (*a*) the frequency of first occurrence as figure, at the beginning of each trial; (*b*) the average duration of the figural period (a figural period designates the time in seconds from the appearance of the plus-figure to the disappearance of the cross-figure and it is equal to the sum of both alternating figures: $t_f = t_p + t_c$); (*c*) the average duration of alternating (plus- and cross-) figures; (*d*) the relative dominance of the plus-figure (t_p/t_c) and the cross-figure (t_c/t_p); and (*e*) the coefficient of variation ($V = 100\ SD/\text{Mean}$) of the figural period and its components (plus- and cross-figures).

Frequency of first occurrence.—The dominant figure tends to be seen first

TABLE 1

FREQUENCY OF FIRST OCCURRENCE OF PLUS-FIGURE

$N = 88$

Angle of Plus-Figure	Obtained		Corrected	Theoretical
	F	%		
0°	—	—	—	90.0
5°	87	98.9	87.5	89.4
15°	86	97.7	83.5	84.6
25°	85	96.6	75.6	75.7
35°	83	94.3	61.3	63.7
45°	77	87.5	50.0	50.0
55°	63	71.6	38.7	36.3
65°	40	45.5	24.4	24.3
75°	27	30.7	16.5	15.4
85°	21	23.9	12.5	10.6
90°	—	—	—	10.0

when a reversible pattern is observed. The frequencies of first occurrence of the plus-figure are given in Table 1. The values for the cross-figure are either 88 or 100% minus those given for the plus-figure. It may be seen that the magnitude of angle greatly influences the frequency with which one of the alternating figures was seen first. The small angles clearly dominated. The curves in Fig. 2 show that each of the alternative plus- and cross-figures has the greatest dominance at its smallest intrafigural angle. The more the angle increases, the more the dominance decreases. In other words, the less the difference between the angles of alternative figures in a reversible pattern, the more the dominance of one of the figures diminishes. The dominance of the plus-figure, however, goes over to the dominance of the cross-figure between 55° and 65°, not at 45°.

Table 1 and Fig. 2 show that there is a preference for the "plus"-figure, which may be explained by the influence of the vertical-horizontal position (plus) as compared with the tilted or oblique direction (cross). In order to get the values without this preference, we may compute these for the total group by formula

$$v = \frac{a - b}{2}$$

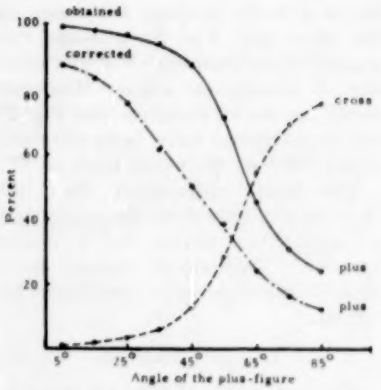

FIG. 2. Frequency of first occurrence as figure. The upper curve of the plus-figure marked "obtained" represents the values obtained from experiments, the lower one marked "corrected" represents the same values corrected for the influence of the vertical-horizontal position. The curve of the cross-figure represents the values obtained from experiments.

TABLE 2

DURATION OF THE FIGURAL PERIOD IN SECONDS AT DIFFERENT ANGLES OF THE PLUS-FIGURE

$N = 60$ (In parentheses N is less than 60)

0	Angle of the Plus-Figure									Mean	*SD*
	5°	15°	25°	35°	45°	55°	65°	75°	85°		
1	4.78	4.46	3.95	2.63	2.65	2.88	4.48	3.90	4.63	3.82	.82
2	4.91	3.79	3.31	3.32	3.13	3.70	3.55	4.32	4.46	3.83	.57
3	3.42	2.93	2.78	3.28	2.74	2.74	2.59	2.77	2.86	2.90	.26
4	2.89	2.99	2.56	2.64	2.86	2.67	2.89	2.94	2.80	2.80	.14
5	3.14	2.74	2.99	2.78	2.69	2.65	2.87	3.00	2.90	2.86	.15
6	(10.75)	3.58	2.76	3.14	2.53	2.52	2.85	2.82	5.24	4.02	2.52
7	6.64	4.18	3.67	3.44	2.88	3.10	3.94	4.52	5.88	4.25	1.19
8	3.21	2.63	2.26	2.18	2.01	2.33	2.38	2.68	3.15	2.54	.40
9	(11.10)	4.13	3.65	3.47	3.90	3.68	4.06	4.67	7.09	5.08	2.37
10	(12.92)	5.58	3.36	3.54	3.30	3.32	3.18	3.95	4.59	4.86	2.94
11	6.23	3.94	3.75	2.98	2.79	2.60	2.76	2.96	3.70	3.52	1.06
Mean	6.36	3.72	3.19	3.04	2.86	2.93	3.23	3.50	4.30	3.68	
SD	3.44	.84	.52	.41	.45	.44	.65	.74	1.33		

where a = the average frequency of the plus- and b = the average frequency of the cross-figure at the same size of intrafigural angle. The corrected curves for the plus- (see Fig. 2) and cross-figures have been obtained which differ by the phase angle of 90°.

The figural dominance does not vary rectilinearly with the angle, but is apparently similar to a cosine function. Theoretical values have been obtained for the plus-figure by formula

$$x_p = a + b \cos 2\phi \qquad (1)$$

and for the cross-figure by formula

$$x_c = a + b \sin (2\phi - 90°) \qquad (2)$$

where $a = 1/2$, $b = 2/5$ and ϕ = intrafigural angle of the plus-figure. The theoretical values for the plus-figure are compared with the empirical data in Table 1. The differences are rather negligible. Consequently, it may be stated that figural dominance, as measured by frequency of first occurrence, is a cosine function of the angle of the sectors which constitute this figure or of the intrafigural angle.

Duration of the figural period.—The average durations of the figural period in seconds for each *O* at different angles of the plus-figure are presented in Table 2. Although the individual means of any column are rather different, the rows of the table show a very distinct general tendency. Shifting of the periods is slowest at 5° for the plus-figure, and quickest at about 45°.

The curve in Fig. 3 shows this general tendency for the total group of *O*s. The duration of figural periods is greatest at the smallest intrafigural angle of the plus-figure where the dominance of this figure is greatest. With the increase of the angle the shifting of the periods accelerates, most quickly between 5° and 15°.

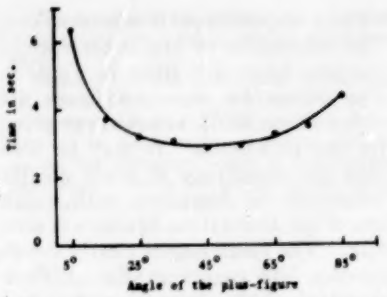

FIG. 3. Average duration of the figural period.

About 45° the periods are shortest. After 65° the shifting of the periods begins to retard again. At 85°, where the intrafigural angle of the cross-figure is smallest (5°) and the dominance of that figure greatest, the shifting of the figural periods has been slowest again, but at that time not so slow as at 5° for the plus-figure. This difference of the duration of figural periods at 5° and 85° has been increased by the results of *O*s numbers 6, 9, and 10 who because of the extremely great dominance of the plus-figure were at 5° apparently at the limit of the reversals and therefore had particularly slow shiftings. But it did not explain this difference entirely because subtracting these *O*s there remains at 5° 4.40 sec. and at 85° 3.80 sec. which is no small difference. It may be, again, explained by the influence of the vertical-horizontal position of the plus-figure.

Duration of the alternating plus- and cross-figures in a figural period.—A figural period consists of two part-figures: plus and cross. These alternating figures may have very different relative duration. Table 3 gives the average durations of the plus-figure in seconds for each *O* at different angles of this figure. The values for the cross-figure are the corresponding values for the figural period minus

TABLE 3

DURATION OF THE PLUS-FIGURE IN SECONDS AT DIFFERENT INTRAFIGURAL ANGLES

$N = 60$ (In parentheses N is less than 60)

O	Angle of the Plus-Figure									Mean	*SD*
	5°	15°	25°	35°	45°	55°	65°	75°	85°		
1	3.47	3.18	2.77	1.52	1.52	1.46	1.97	1.77	1.92	2.18	.72
2	2.71	2.08	1.67	1.68	1.60	1.97	1.74	2.17	2.24	1.98	.34
3	1.96	1.59	1.51	1.67	1.45	1.38	1.25	1.36	1.48	1.52	.20
4	2.02	1.82	1.39	1.40	1.46	1.34	1.46	1.46	1.38	1.53	.22
5	1.95	1.63	1.67	1.44	1.39	1.34	1.49	1.43	1.33	1.52	.19
6	(9.11)	2.32	1.54	1.81	1.35	1.31	1.42	1.24	1.45	2.39	2.60
7	3.63	2.04	1.79	1.66	1.44	1.52	1.90	1.98	1.73	1.97	.62
8	2.40	1.87	1.43	1.30	1.09	1.23	.97	1.00	1.11	1.38	.44
9	(7.75)	3.08	2.29	1.90	2.38	2.14	2.13	2.15	3.49	3.03	1.74
10	(10.97)	3.42	1.77	1.88	1.59	1.59	1.39	2.13	2.18	2.99	2.88
11	5.52	2.98	2.17	1.73	1.45	1.28	1.23	1.02	1.05	2.05	1.36
Mean	4.68	2.36	1.82	1.64	1.52	1.51	1.54	1.61	1.76	2.05	
SD	3.05	.64	.41	.19	.30	.28	.34	.43	.69		

those given for the plus-figure. Although the individual differences are very great the general tendency is the same. The smallest angle of 5° has the greatest duration. Figure 4 shows the average duration of the plus-figure at different angles for the total group of *O*s as compared with those of the cross-figure. It may be seen that the dominance of two alternating figures varies inversely with the intrafigural angle. The duration of the plus-figure is greater until 45°, between 45° and 65° both figures have nearly the same duration, but above 65° the cross-figure has the dominance.

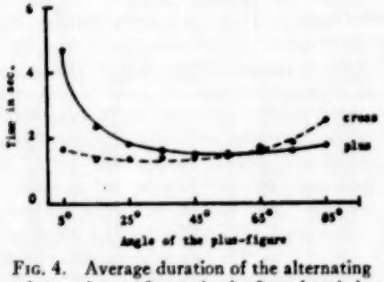

FIG. 4. Average duration of the alternating plus- and cross-figures in the figural period.

Relative dominance of the plus- and cross-figures.—Relative dominance of the plus-figure means the relation of the duration of the plus-figure to that of the cross-figure (t_p/t_c) and relative dominance of the cross-figure its inverted value (t_c/t_p). The values of Table 4 and the curves as plotted in Fig. 5 show very clearly that the relative dominance of both alternating figures increases with the decrease of the intrafigural angle, and vice versa,

TABLE 4

RELATIVE DOMINANCE OF THE PLUS- (t_p) AND CROSS-FIGURE (t_c) IN THE FIGURAL PERIOD

$(t_p + t_c)$

$N = 660$

Plus-Figure		Cross-Figure	
Angle	t_p/t_c	Angle	t_c/t_p
5°	2.79	85°	.36
15°	1.74	75°	.58
25°	1.33	65°	.75
35°	1.17	55°	.85
45°	1.13	45°	.88
55°	1.06	35°	.94
65°	.91	25°	1.10
75°	.85	15°	1.17
85°	.69	5°	1.44

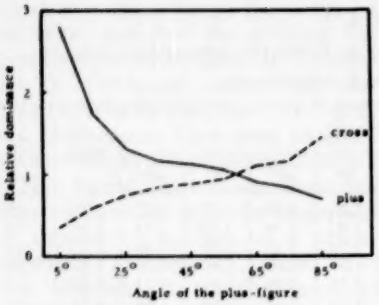

FIG. 5. Relative dominance of the plus- and cross-figures.

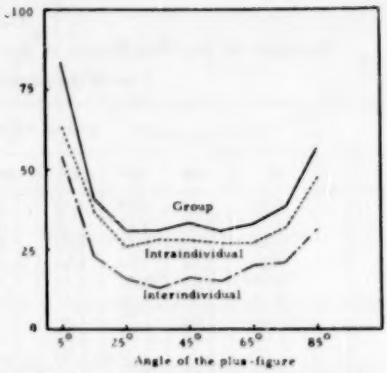

FIG. 6. Variation of the figural period.

and that the dominance of the plus-figure is greater than that of the cross-figure. Both curves intersect between 55° and 65° which means that the indifferent nondominant zone is about here. The same nondominant zone was found also at the first occurrence as measure of figural dominance.

Coefficient of variation of the figural period and its components.—The coefficient of variation (V) of the whole figural period, as well as plus- and cross-figure components, may also be used as one measure of figural dominance, too. Table 5 gives Vs of the intraindividual, interindividual and group variations for the figural period and its components. It is easily seen that in all cases V has its highest value at the smallest angle in the vertical-horizontal position. Also at the smallest intrafigural angle of the cross-figure (or at 85° of the plus-figure) the Vs are higher than in the middle range of angles. This means that the reversals are more regular in the middle range of angles than at the extremes. Comparing the values of the figural period with those of its components it may be stated that the figural period is more regular than the plus- and cross-figures.

Values of Vs as plotted in Figure 6 show that the intraindividual variation is greater than the interindividual. The same tendency appears also in the plus- and cross-figures. So the former seems to influence more the variation of the group than the latter.

TABLE 5

COEFFICIENTS OF VARIATION (V) OF THE FIGURAL PERIOD AND ITS COMPONENTS

Angle of Plus-Figure	Figural Period			Plus-Figure			Cross-Figure		
	Intra-indiv.	Inter-indiv.	Group	Intra-indiv.	Inter-indiv.	Group	Intra-indiv.	Inter-indiv.	Group
5°	64	54	84	70	65	96	69	50	80
15°	37	23	41	44	27	52	37	32	49
25°	26	16	31	35	23	41	31	20	36
35°	28	13	31	34	12	35	37	19	41
45°	28	16	33	34	20	39	31	16	34
55°	27	15	31	32	19	38	31	13	34
65°	27	20	33	35	22	42	36	21	41
75°	32	21	38	37	27	46	38	20	43
85°	47	31	56	55	39	66	45	36	58

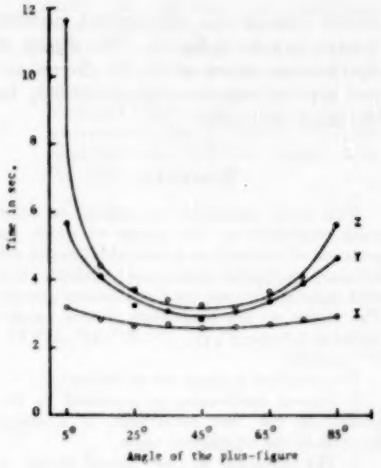

FIG. 7. Average duration of the figural period for Types X, Y, and Z.

Individual differences.—From Table 2 it may be seen that individual differences in the average duration of the figural periods are very great. The main difference between *O*s is the degree of variation in the duration of the figural periods at different angles. One *O* may have approximately equal duration for his figural periods at different angles, while those of another *O* vary greatly with angles. In this regard there may be distinguished three types of curves which shall be designated as X, Y, and Z (see Fig. 7). O_3, O_4, O_5 and O_8 belong to Type X; O_1, O_2, O_7 and O_{11} to Type Y; O_6, O_9 and O_{10} to Type Z.

Type X has nearly the same times at all angles. Figural periods are rapid, shiftings of the part-figures frequent.

Type Z is opposite to X. The duration of the figural periods varies greatly. Shiftings of the part-figures depend to a great degree on the size of the intrafigural angle. They are generally slow. At 5° the dominance of the plus-figure is already so great that the limit of reversals seems to be reached. Reversals could be perceived here only sometimes, sometimes not at all.

Type Y lies between the two extreme types. The *O*s belonging here are influenced by the extreme angles, but not to the same degree as Type Z.

Although Type X has nearly the same duration for the whole figural period at different angles, it may be seen from Fig. 8 that this is not the case for the plus- and cross-figure components. They both have their longest duration at the smallest intrafigural angle and decrease with the increasing angle. The decrease of the average durations with the increasing angle is greater for Type Y and greatest for Type Z (see Fig. 8).

Phenomenal observations.—The alternation of reversible figures was rather irregular. The rate of shifting was highly variable. Average rates of the total group ranged from 15 to 20 per minute and agreed with those of Hollingworth (**4**) and Tussing (**10**). But we could not find any particular trend in the durations of alternating figures. Thus, we could not confirm the increasing rate of the alternation of two part-figures which were reported by Köhler (**6**). Among the 11 *O*s there were only two who could report that they often had an

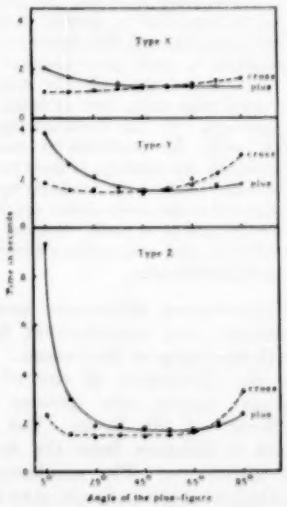

FIG. 8. Average duration of the alternating plus- and cross-figures for Type X, Y, and Z.

accelerating rate of alternations. Also Oyama and Torii (7) reported that the prolonged observation did not cause the acceleration of the rate of reversal.

In no less degree variable is the clearness of the alternating figures. All *O*s reported variations in clearness during the continued observation.

We found that the dominant figure is much more "figural" than the nondominant. The phenomenal differences between the dominant and nondominant figures may be summarized as follows: (*a*) The dominant figure has the tendency to dominate the figural field. When the nondominant figure appears as figure, it does not have that tendency. (*b*) The dominant figure remains in figural position for a long time. The nondominant can be seen as figure only for a short time; sometimes it only "glimpses" as the *O*s say. (*c*) The dominant figure appears as figure very rapidly, sometimes so suddenly that *O* cannot report it at the right moment. The nondominant figure appears as figure more slowly. (*d*) The dominant figure, as figure, has always much greater clearness than the nondominant in the same role. (*e*) The dominant figure appears as figure always entirely, as a whole, while the nondominant can lose one or another of its part. (*f*) The contours of the dominant figure are noticeably sharper than those of the nondominant. (*g*) The dominant figure has the tendency to be seen first as figure, while the nondominant in general does not appear first as figure. (*h*) The clearness of the dominant figure is much more constant than that of the nondominant. (*i*) The dominant figure is much more easily seen as figure than the nondominant. (*j*) The dominant figure is aggressive, while the nondominant may be characterized, on the contrary, as more passive. (*k*) In the cases of fatigue the dominant figure is more resistant than the nondominant which may loose one or another of its parts. Some of these points have been used as quantitative measures for the figural dominance.

The phenomenal differences between the dominant and nondominant figure vary with the degree of dominance. The greater the dominance of one of the alternating figures, the greater the "figuredness" of this figure. But this difference is different from the figure-ground distinction. The nondominant figure, when it is seen as figure, may have very low figuredness, but it retains its figural quality and will at that moment always be seen as figure. The degree of figuredness varies with the dominance and may be expressed quantitatively by the same measures.

Summary

This study represents an attempt to determine quantitatively the extent to which the principle of proximity in a reversible pattern will influence the figural dominance. Rubin's circle with eight sectors was used as stimulus pattern. The angles of the plus-figure of this pattern varied as follows: 5°, 15°, 25°, 35°, 45°, 55°, 65°, 75°, and 85°.

The principal findings are as follows:

1. Figural dominance, as measured by frequency of the first occurrence, is a cosine function of the intrafigural angle.

2. The duration of the figural period is dependent upon the intrafigural angles of the alternating figures. It is shortest at 45°. The greater the difference between the intrafigural angles of these figures, the longer the duration of the figural period.

3. The duration of the part-figures in a figural period shows that the dominance of the alternating figures varies inversely with the intrafigural angle.

4. The individual differences are very great.

5. Great phenomenal differences are found between the dominant and nondominant figures. The former is much more "figured" than the latter. The degree of figuredness of one of alternating figures varies directly with the dominance of that figure.

References

1. Goldhamer, H. The influence of area, position, and brightness in the visual perception of a reversible configuration. *Amer. J. Psychol.*, 1934, **46**, 189–206.
2. Graham, C. H. Area, color, and brightness difference in a reversible configuration. *J. gen. Psychol.*, 1929, **2**, 470–483.
3. Harrower, M. R. Some factors determining figure-ground articulation. *Brit. J. Psychol.*, 1936, **26**, 407–424.
4. Hollingworth, H. L. Perceptual fluctuation as a fatigue index. *J. exp. Psychol.*, 1939, **24**, 511–519.
5. Koffka, K. *Principles of Gestalt psychology.* London: Routledge and Kegan Paul, 1935.

6. KÖHLER, W. *Dynamics in psychology.* New York: Liveright, 1940.
7. OYAMA, T., & TORII, S. Experimental studies of figure-ground reversal: I. The effects of area, voluntary control and prolonged observation in the continuous presentation. *Jap. J. Psychol.*, 1955, **26,** (Japanese text, 178–188; English text, 217–218).
8. RUBIN, E. *Synsopleveде Figurer.* Copenhagen: Gyldendalske, 1915.
9. RUBIN, E. *Visuell wahrgenommene Figuren.* Copenhagen: Gyldendalske, 1921.
10. TUSSING, L. Perceptual fluctuation of illusions as a possible physical fatigue index. *J. exp. Psychol.*, 1941, **29,** 85–88.

(Received December 27, 1955)

Journal of Experimental Psychology
Vol. 53, No. 1, 1957

REINFORCEMENT OF VERBAL BEHAVIOR [1]

DOUGLAS M. McNAIR [2]

University of North Carolina

The experiment reported here represents an attempt to explore some relations between linguistic behavior and independent variables without assuming that verbal reactions differ fundamentally from other forms of behavior. This was done in the belief that the important psychological problems of language concern the establishment of lawful relations between verbal behavior, independent variables, and interactions of independent variables. The experiment was conducted and interpreted within the framework of Skinner's system (**9, 10, 11**). Primarily, interest is in some functional relations between speech as a dependent variable and reinforcement and referents of speech as independent variables.

A number of recent experimental manipulations of verbal behavior within the operant conditioning framework have been reported. Greenspoon (**2**) used a free responding procedure in which *S*s spoke any single words occurring to them. Reinforcement was contingent upon specific response classes. He found that one verbal stimulus ("mmm-hmm") and two nonverbal stimuli (a tone and a light) were reinforcing for members of a relatively small, rigidly defined response class (plural nouns). For a larger, heterogeneous class of responses (all words except plural nouns), two verbal and two nonverbal stimuli ("huh-uh" and the three stimuli above) were reinforcing.

Taffel (**12**) also reported that the reinforcement operation increased the frequency of a small class of responses (first person pronouns). Cohen, Kalish, Thurston, and Cohen (**1**), using Taffel's procedure, confirmed Taffel's results. In addition they found that by shifting reinforcement from one class of responses to another they could alter the response pattern. Thus, when reinforcement was shifted from first to third person pronouns, there was a consequent shift in response frequencies.

In a paper published after the completion of the present study, Sidowski (**8**) has indirectly anticipated some results reported here. It seems apparent from the slopes of Sidowski's curves that rate of verbal response is a function of rate of reinforcement, although he did not measure rate of response per se nor did he systematically vary rate of reinforcement.

The foregoing studies have offered evidence supporting theoretical extension of the reinforcement principle, based largely on experimentation with infrahuman *S*s, to verbal behavior. In addition they have demonstrated that reinforcement contingent upon clearly defined classes of verbal responses probably influences the rate of emission of such responses.

The present experiment is an endeavor to extend the operant conditioning paradigm to verbal behavior in a situation resembling the speaker phase of a typical conversational

[1] This paper is a portion of a dissertation submitted to the Department of Psychology of the University of North Carolina in partial fulfillment of the requirements for the Ph.D. degree. The writer is indebted to Dr. Irvin S. Wolf for invaluable advice and guidance throughout the course of the study and to Dr. B. J. Winer for help with the statistical analysis.

[2] Now at The Woman's College of the University of North Carolina.

event (**4, 7**). The study concerned the influence of reinforcement upon total amount of talk about certain topics. Since frequencies of words within fixed time intervals were observed as the dependent variable, rates of verbal responding were the primary concern. The major hypotheses tested were: (*a*) Rates of verbal responding to a referent may be conditioned by operant reinforcement; (*b*) different rates of reinforcement have distinguishable effects upon verbal response rates.

Method

Apparatus.—The apparatus consisted of a 35-mm. projector, a wall screen, four 35-mm. slides, electrodes leading to a complex-appearing panel of dials and switches, a single-tone door chime with a silent pushbutton, a tape recorder and a stopwatch. The four 35-mm. slides used as referential stimuli were: (*a*) Practice slide: a horse in coronation garb; (*b*) Slide P_2: a family gathered before a fireplace; (*c*) Slide P_3: a camp counselor comforting a distraught camper; (*d*) Slide P_4: a man and woman embracing. Photographic reproductions of these slides are presented elsewhere (**6**).

Experimental plan.—The *S*s were 54 volunteers from undergraduate psychology courses. Eighteen *S*s were assigned (by restricted randomization) to each of three groups. All *S*s talked for 2-min. periods about each of the three interpersonal pictures. Group R_8 received a "high" rate of reinforcement while they talked. The rate was "low" for Group R_2. The reinforcing stimulus (RS) was not presented for Group R_0.

Presentation of the RS was not dependent upon the occurrence of any particular word or class of words. It was introduced at moments in time which were predetermined. The only requisite for the presentation of the RS was that *S*s had to be verbalizing at the moment. The RS was presented for Groups R_2 and R_8 by the variable-interval method. A bell-tone signifying approval was presented once in each 15-sec. interval for *S*s in Group R_8. It thus occurred eight times in association with each picture. The same RS was presented once in each 60-sec. interval (twice with each picture) for *S*s in Group R_2. The RS was presented in each 15- or 60-sec. interval at a moment previously determined by randomization. In the event that *S*s were silent at the selected instant, the RS was presented immediately after they resumed talking.

The rationale for selecting the RS requires some elaboration. The RS (bell-tone) employed in the experiment was defined in instructions to the *S*s as signifying approval. Theoretically it appears to qualify as one of a large number of secondary reinforcers and was selected on this basis. Until tested under the experimental conditions, however, it could not be accurately described as a reinforcer, if one follows Skinner's concept of reinforcement as a stimulus following a response which alters the probability of that response (**11**, pp. 72–73).

While the RS was defined to *S*s as an indicator of approval, the relationship of the RS to their behavior was disguised. They were instructed that the approval was related to one aspect of behavior—tension reduction. (Any real relation to this component was chance.) Actually the RS was correlated with talk about the pictures. This procedure was adopted primarily to prevent *S*s from recognizing the purpose of the experiment and to obviate attempts by *S*s to form hypotheses about the significance of the RS. In previous research (**1, 2, 12**), the RSs have been introduced without explanation to the *S*s. It is possible, in such a procedure, that an important variable is not only the contingency between a specific response and reinforcement but also the fact that the introduction of an unexplained stimulus creates a problem-solving situation (**3**).

The order of presentation of the three pictures was counterbalanced to isolate the effects of this variable. Three *S*s in each group were assigned (by restricted randomization) to each of the six possible orders. The experimental sequences of slides were: (*a*) Order 1: P_2, P_3, P_4; (*b*) Order 2: P_2, P_4, P_3; (*c*) Order 3: P_3, P_2, P_4; (*d*) Order 4: P_3, P_4, P_2; (*e*) Order 5: P_4, P_2, P_3; (*f*) Order 6: P_4, P_3, P_2.

Procedure.—The *S*s sat in a darkened room approximately 6 ft. from a wall screen. A microphone was placed about 18 in. to the left-front of *S* at face level. Electrodes were attached to the right hand. All other apparatus was on *E*'s desk behind *S*.

Upon entering the room *S*s read instructions describing their task and defining the RS. The manner in which the RS was structured can best be understood from the directions: "This is a study of stress or tension during normal speech. Pictures will be shown to give you something to talk about. All you have to do is talk about them in a normal voice while they are on the screen. As you talk you should show less tension. This is good—and each time your tension drops a bell will sound to let you know.

Read this again. You may ask questions now, but say nothing to the experimenter from the time we begin until we are completely finished."

After the instructions, the practice slide was projected for approximately 1 min. During this period, *Ss* were encouraged to say anything that occurred to them and to continue speaking until the picture was removed. The RS was demonstrated twice following the practice picture. The three experimental slides were then presented for 2-min. periods with intervals of approximately 15-sec. between pictures.

After the experiment *Ss* were questioned about what they considered the purpose of the experiment, the significance of the bell-tone and their reasons for finding it easy or difficult to talk about the individual pictures.

Measurement.—Frequencies of total verbal responses for each *S* for each slide were tabulated directly from the tape recordings. The single word was considered the unit of response. In case of doubt, *Webster's* (**13**) was consulted as the authority for what constitutes a single word. In addition these rules were followed: contractions were counted as the number of words contracted; abbreviations were counted as the number of words abbreviated; interjections were not counted.

To secure a measure of the reliability of the word-counts, a random sample of 25 2-min. response periods was submitted to a second individual. The reliability coefficient for the word-counts of the two judges was .99.

Results

To compare the effects of the experimental variables upon total response rates of the three groups, an analysis of variance technique for repeated measurements upon the same *Ss* was employed. The variables in this analysis were: groups (reinforcements), orders, pictures, and positions (first, second, third) within the experimental sequence of pictures. Since the order of presentation of the pictures was counterbalanced, the six possible permutations represent two 3 × 3 Latin squares with Latin letters for positions. This analysis permitted not only a test of differences in response rates as a function of the RS. It also permitted the isolation of variation due to the other experimental variables and their interactions.

To test differences in response rates between groups (reinforcement rates), between orders and the interaction of reinforcement × order, an error term based on independent, randomly assigned *Ss* was used. This was the mean square based on variation between *Ss* within groups and orders. In order to test the differences in response rates as a function of differences in stimulus value of the three pictures and as a function of position within the experimental sequence of pictures, another error term was used. This second error term was also used to test interactions of pictures and positions with the other experimental variables. The error term used in these comparisons was the mean square based on the pooled residual variation within *Ss*. Possible correlation is thus taken into consideration as these comparisons are based on repeated measurements of the same *Ss*.

The results of the four-way analysis of variance are presented in Table 1. The *F* ratio for differences between

TABLE 1

Analysis of Variance of Verbal Responses to Three Pictures Presented in Six Orders with Three Reinforcement Rates

Source of Variation	*df*	Mean Square	*F*
Reinforcement rates (R)	2	36,395.29	3.94*
Orders (O)	5	29,333.86	3.17*
O × R	10	4,469.81	
Ss/R/O†	36	9,240.98	
Total between *Ss*	53		
Pictures (P)	2	7,063.23	21.61***
P × R	4	190.47	
P × O	10	1,184.27	3.62**
P × R × O	20	446.70	1.37
Positions	2	3,624.03	11.09***
Positions × R	4	731.79	2.24
Positions × O	10	1,872.11	5.78***
Positions × R × O	20	249.58	
Pooled residual	36	326.87	
Total within *Ss*	108		
Total	161		

* $P < .05$.
** $P < .01$.
*** $P < .001$.
† Further analyzed into variation between *Ss* within each rate, Bartlett's test applied, and no basis found for rejection of hypothesis of homogeneity of variance.

groups receiving different rates of reinforcement permits rejection, with reasonable confidence, of the null hypothesis that differences between groups arose from a random sampling of a common population. The inference that the RS reinforced verbal-response rates is supported by the data. We also have evidence that the two reinforcement rates differentially affected response rates. The higher rate of reinforcement had the greater facilitating effect. (Mean verbal responses for rates 0, 2, and 8 were 568.7, 635.1, and 723.9. Tukey significant gap test for $t_{.05}$ with 36 *df* yields critical difference of 64.9.)

The 2-min. response periods for each picture were divided into four 30-sec. intervals. This was to focus upon trends in response rates as a function of intervals within the total time of exposure of each picture. Table 2 presents mean-response rates for each group during each 30-sec. interval. Data for all three pictures were combined in obtaining these means. For all groups there is a decline in response rates as the time of exposure increased. To compare these trends, a separate analysis of the data classified by reinforcements and intervals was performed. If the difference in trends for the groups is significant, this would be indicated by a significant Reinforcement × Interval interaction (**5**). The complete two-way analysis of variance is not reported, but the Reinforcement × Interval interaction yielded an F of 2.98 (for 6 and 153 *df*, $P < .01$). This would indicate that the hypothesis that these trends do not differ can be rejected with considerable confidence. (This analysis is admittedly crude since a more complex one would permit removal of variation due to pictures, positions, and various interactions which inflate the error term, but it should be sufficient for this purpose.) Response rates for Group R_8 remain high relative to the other groups, and the influence of the RS seems most evident during the second minute of exposure. Means for all groups are near the same level for the first 30 sec. Rates for Group R_0 decline rather rapidly compared to R_8 and R_2, and the downward trend in responsiveness for R_2 is intermediate between the high and zero reinforcement groups.

TABLE 2

MEAN VERBAL RESPONSES DURING 30-SEC. INTERVALS OF EXPOSURE OF PICTURES AS DETERMINED BY REINFORCEMENT RATES (0, 2, 8)

Group	30-Sec. Interval			
	1st	2nd	3rd	4th
R_8	65	61	58	58
R_2	59	54	51	47
R_0	60	49	44	38
Mean	61	55	51	48

The sequences in which the pictures were presented apparently influenced verbal response rates (Table 1). Tukey significant gap and "straggler" tests were employed for comparing individual verbal response means for the six orders. These tests separated the orders into three groups: Orders 1 (M = 788.8) and 3 (M = 719.8); Orders 2 (M = 649.0), 6 (M = 593.6) and 4 (M = 589.3); and Order 5 (M = 515.0).

Orders 1 and 2 began with the same picture (P_2). This was also true for Orders 3 and 4 (P_3) and Orders 5 and 6 (P_4). Data for orders beginning with the same picture were combined, and the Tukey gap test was employed to compare individual means. This test separated combined Orders 5 and 6 (M = 554.3) from Orders 1 and 2 (M = 718.9) and Orders 3 and 4 (M = 654.6). (Significant gap = 64.9, $t_{.05}$, 36 *df*.) Beginning the sequence of slides with P_4 depressed the rate of response for the entire sequence of pictures.

Slide P_4 not only lowered response rates when it was the first picture in a series, but it seemed to have a relatively low stimulus value for verbal responses. Differences in verbal-response

rates attributable to differences between pictures is significant (Table 1). The significant gap test separates P_4 (M = 201.1) from P_2 (M = 222.1) and P_3 (M = 219.4). (Significant gap = 7.1, $t_{.05}$, 36 *df*.)

Rate of response differences attributable to the various positions of the experimental sequence (independent of specific picture) are significant (Table 1). The gap test for comparing individual means separates the first position (M = 205.1) from the second (M = 216.5) and third (M = 221.0) positions. (Significant gap = 7.1, $t_{.05}$, 36 *df*.)

Among the interactions, primary interest was in possible interactions between reinforcements and other variables. Within the range of this experiment, effectiveness of reinforcement does not appear to vary significantly as a function of specific topic of speech, sequence of a series of referents, or position within a series of referents (Table 1). Two significant interactions, Pictures × Orders and Positions × Orders, were noted (Table 1). Apparently, the influence of a specific topic of speech on rate of response depends on the sequence in which it is presented. Also, the effect of a specific position within an order varies with the sequence involved.

Only two *S*s verbalized any suspicion of deception. No *S* related amount of talk to the RS. Most *S*s in Group R_0 admitted concern about the absence of RS. Approximately half the *S*s in Groups R_2 and R_3 stated they "paid some attention" to the RS. The other half said, in effect, they ignored the tone. Nearly all *S*s who expressed any concern about RS reported this occurred on the first picture and that they soon "forgot" about it. A few *S*s in R_2 and R_3 reported feeling decreased tension at approximately the times RS was presented. Some *S*s in R_0 "rationalized" their failures to register decreased tension, stating they had run to be on time for the experiment, etc. A few *S*s in both R_0 and R_2 complained of running out of words, adding that this was unusual for them. No remarkable response patterns were associated with these verbal reports of *S*s.

Discussion

From the results, it appears that amount of talk about a topic within a given time interval (rate) is influenced by the reinforcement operation. And, at least within limits, differential rates of reinforcement have distinguishable effects upon rates of verbal responding. Apparently the rate variable is quite effective since it was found to be statistically significant with only three levels. A more exhaustive study of the rate dimension would contribute more but was beyond the scope of this study (6). The importance of finding rate effective in controlling verbal behavior would be enhanced if extended and refined in subsequent research to apply to the individual, i.e., where different rates of reinforcement are applied to single *S*s. This was attempted, with inconclusive results, in a preliminary experiment (6). A related issue would be the determination of what operations are reinforcing and to what extent. Presumably some sort of hierarchy of secondary reinforcers could be established which could be effectively utilized to increase the control of such events as psychotherapy and language learning.

The effectiveness of the RS in this experiment appeared most obvious in later segments of the exposure period of the slides. It appeared as if a number of statements were rather easily produced in relation to each picture. After this period, *S*'s verbal behavior may have become more a function of the RS. It is possible that reinforcement administered only during this second-minute period would have been as effective as reinforcement throughout. This question may have practical implications. In many interview situations it would be important to know if reinforcement is most effective during periods when the

individual has difficulty continuing with a topic. A related question would be: is reinforcement as effective, or does it become ineffective, when presented indiscriminately at times when other variables are operating to maintain responsiveness?

Certain aspects of the referential stimuli evidently exerted an influence upon total response rates. Rate of verbal response is a function not only of the specific subject one talks about but also of the sequence in which a series of topics is presented and the position of a referent within the sequence. Of the three experimental slides, the picture of the man and woman embracing (P_4) produced relatively low response rates. Apparently this picture either functioned to evoke responses incompatible with verbalization or *S*s had prior histories of lesser reinforcement for talk concerning this kind of situation. The effects of the position within a sequence may reflect the operation of adaptation or practice, since the rate of response increased after the first position.

Beginning the series of slides with the picture of the man and woman embracing had a depressing effect upon the whole series, at least as far as rate of verbal response is concerned. This suggests that not only present but preceding referential stimuli affect speech rates. This interpretation is supported by the Pictures × Orders interaction. Implications from this finding might be relevant to many areas. One could argue that, if amount of information is a function of amount of talk, the sequence in which topics are broached would be an important variable in many interview situations. One would also question the effects of sequence on productivity to many tests, especially projective techniques. Obviously, experimentation would be necessary to isolate the specific factors responsible for such sequence effects. The present results suggest that beginning with a somewhat "emotional" topic depresses total or subsequent productivity.

From the fact that *S*s seemed generally unaware of the deception employed in the study, it appears as if reinforcement may act as a controlling variable even when *S*s do not know to which aspect of their behavior it is related. This result is generally in accord with the results of others (**1, 2, 8, 12**). Unlike these other procedures, the relation of RS to the reinforced behavior was actively disguised in this study. It is believed that the present method minimized tendencies of *S*s to "cooperate" to produce results or to structure the situation as a problem to solve. If true, this would seem to be more typical of the usual conditioning experiment.

Our findings lend further support to the attempt by Skinner (**10**) to extend the operant conditioning paradigm to verbal behavior. In this respect, our results substantiate other findings (**1, 2, 12**).

Summary

Three groups of 18 *S*s received zero, "low," and "high" rates of reinforcement while talking about three photographic slides depicting interpersonal situations. The RS was presented at variable intervals. Its relationship to the verbal responses of the *S*s was disguised. Total verbal responses of each *S* to all slides were recorded and analyzed as the dependent variable.

The results support the following major conclusions: (*a*) The reinforcement operation influenced rates of responding to the referents. (*b*) Differential rates of reinforcement had distinct effects upon rates of responding. The higher rate of reinforcement had a greater facilitating effect. (*c*) The reinforcement operation may be effective when *S* is not aware of the contingency between RS and the behavior to which it is related. (*d*) Amount of talk about the same referential stimuli varies as a function of the sequence and the position within the sequence in which they are presented.

References

1. Cohen, B. D., Kalish, H. I., Thurston, J. R., & Cohen, E. Experimental manipulation of verbal behavior. *J. exp. Psychol.*, 1954, **47**, 106–110.
2. Greenspoon, J. The effect of verbal and non-verbal stimuli on the frequency of members of two verbal response classes.

Unpublished doctor's dissertation, Indiana Univer., 1951.

3. Irwin, F. W., Kaufman, K., Prior, P., & Weaver, H. B. On "learning without awareness of what is being learned." *J. exp. Psychol.*, 1934, **17**, 823–827.
4. Kantor, J. R. *Principles of psychology.* Bloomington, Ind.: Principia Press, 1949. Vol. II.
5. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.
6. McNair, D. M. Reinforcement of verbal behavior. Unpublished doctor's dissertation, Univer. of North Carolina, 1954.
7. Pronko, N. H. Language and psycholinguistics: a review. *Psychol. Bull.*, 1946, **43**, 189–239.
8. Sidowski, J. B. Influence of awareness of reinforcement on verbal conditioning. *J. exp. Psychol.*, 1954, **48**, 355–360.
9. Skinner, B. F. *The behavior of organisms.* New York: Appleton-Century, 1938.
10. Skinner, B. F. *Verbal behavior.* William James Lectures, Harvard Univer., 1948.
11. Skinner, B. F. *Science and human behavior.* New York: Macmillan, 1953.
12. Taffel, C. Conditioning of verbal behavior in an institutionalized population and its relation to "anxiety level." Unpublished doctor's dissertation, Indiana Univer., 1952.
13. *Webster's new international dictionary of the English language.* (2nd Ed., Unabridged) Springfield, Mass.: Merriam, 1934.

(Received November 21, 1955)

Journal of Experimental Psychology
Vol. 53, No. 1, 1957

THE EFFORT VARIABLE IN THE ACQUISITION, EXTINCTION, AND SPONTANEOUS RECOVERY OF AN INSTRUMENTAL RESPONSE [1]

EDWIN G. AIKEN

University of Illinois [2]

In 1943 Mowrer and Jones (**5**) performed a bar-pressing experiment to study the influence of the effortfulness of responding on the rate of extinction. They obtained a decrease in trials to extinction with increased effortfulness. Recently, Applezweig (**1**) and Maatsch, Adelman, and Denny (**4**) have pointed out that because Mowrer and Jones employed different numbers of learning trials under their different effort levels, interpretation of the results in terms of Hull's inhibition theory (**3**) is made ambiguous. Applezweig and Maatsch, Adelman, and Denny met this problem in separate experiments by equalizing the number of training trials. The latter failed to obtain any clear variation in extinction rate with effort, while Applezweig confirmed the Mowrer and Jones result. However, his study was confounded by a rather severe selection of *S*s for one of his groups.

The purposes of this investigation are to attempt remedies for some of the design difficulties of earlier experiments and to extend the study of effortfulness to the spontaneous recovery of the extinguished response.

[1] Based upon a thesis submitted in 1954 in partial fulfillment of the requirement for the degree of Doctor of Philosophy at the University of Illinois. The writer wishes to express his gratitude to his thesis advisor, Professor G. Robert Grice, for advice and guidance throughout the research, and to Professors Charles E. Osgood and O. Hobart Mowrer for many helpful suggestions.

[2] The writer is now on active duty in the U. S. Army, stationed at the Psychology Department of the Army Medical Research Laboratory at Fort Knox, Kentucky.

METHOD

Apparatus.—The learning task was a panel-pressing problem similar to that employed by Horenstein (**2**). The apparatus (Fig. 1) has small dimensions in order to restrict the amount of *S*'s initial exploratory behavior. The cup at the end of the arm on the swinging door was a holder for weights used to increase the pressure necessary to swing the door forward and expose the food trough. The door had to be swung forward ½ in. to secure food. A block prevented the door from being swung more than 1 in. The guillotine door, which was manually operated by *E*, was raised at the beginning of a trial and closed at the end of a trial. Latency measures were secured by mounting a microswitch so that it was contacted by the top of the guillotine door when it was raised. This contact started a Standard Electric timer which was stopped when the swinging door was displaced ½ in. or more. When the guillotine door was lowered, the timing mechanism was automatically reset for another trial.

Subjects.—The *S*s were 80 albino rats from the colony at the University of Illinois, which were divided into 8 groups of 10 *S*s. The *S*s varied in age from 70 to 110 days at the start of the experiment, and age as well as sex was equated for all groups.

Experimental design.—The total experimental period covered 14 days. From Day 1 to Day 3, *S* was deprived of food. On Day 3 *S* was given a 30-min. free eating period in a compartment amply supplied with Purina Laboratory Chow. This feeding period was repeated on Days 4, 5, and 6, with *S* returned to the home cage during the intervening periods. On Day 8, following approximately 48-hr. food deprivation, pretraining was initiated.

For pretraining *S* was placed in the starting compartment of the apparatus with the guillotine door down. After 30 sec. the door was raised, exposing the swinging door. As *S* approached the swinging door, it was manually pulled back by *E*, thus exposing the food trough

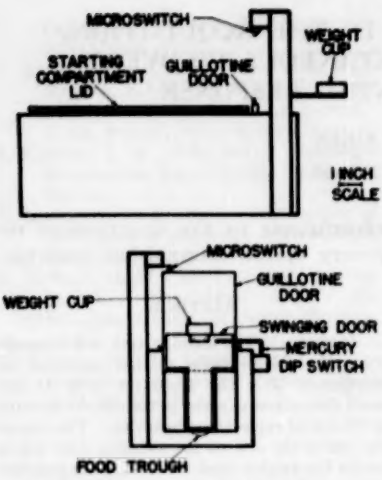

FIG. 1. Experimental apparatus—side and front views.

which was liberally spotted with Purina Chow and one commercially prepared laboratory food pellet weighing 20 mg. After a bit of exploration, *S* would happen upon the filled food trough and begin eating. As soon as the pellet was taken by *S*, *E* lowered the guillotine door, put in a new pellet, and after 30 sec. raised the guillotine door for another pretraining trial. Fifteen such pretraining trials were given with a constant intertrial interval of 30 sec. The *S* was then placed in a feeding compartment for a 30-min. free-eating period before being returned to its home cage.

On Day 10, under 48-hr. food deprivation, *S* was given its first acquisition trials. Thirty seconds after placing *S* in the apparatus, the guillotine door was raised, starting the timer and the first trial. In a short period of time *S* would succeed in swinging the door out far enough to obtain the food pellet and stop the timer. The *E* then closed the guillotine door, recorded the response latency, placed a new pellet in the trough, and after 30 sec., raised the guillotine door to start the second trial. Fifteen such acquisition trials were given on Day 10 with a constant intertrial interval of 30 sec. The *S* was then placed in a free-feeding situation for 30 min. and returned to its home cage. On Days 11 and 12 the procedure was repeated, but now under 24-hr. food deprivation, for a total of 45 acquisition trials.

On Day 13, under 24-hr. food deprivation, the extinction trials were begun. Thirty seconds after placing *S* in the starting compartment of the apparatus, the guillotine door was raised, but now without a food pellet in the trough. After *S* had pushed the door forward an amount sufficient to stop the timer, the guillotine door was lowered, the response latency recorded, and 30 sec. later the guillotine door was raised for the next trial. If *S* failed to respond on any particular trial within 30 sec., the guillotine door was lowered, and a new trial was started 30 sec. later. Experimental extinction was defined as having been reached when *S* failed to respond on two consecutive 30-sec. periods with the swinging door exposed. After reaching criterion, *S* was returned to the feeding compartment for 30 min. and then to its home cage.

On Day 14, again under 24-hr. food deprivation, *S* was tested for the spontaneous recovery of the response. Recovery was indexed by the number of trials *S* took to reach extinction a second time. The operations for Day 14 were the same as for Day 13.

The amount of effort necessary to displace the swinging door was varied by either placing a 50-gm. weight in the weight cup or leaving it empty. A pressure of 5 gm. was necessary to displace the door when empty and a pressure of 32 gm. was necessary when the 50 gm. weight was in the cup.

During the learning trials, the 80 *S*s were divided into two groups: one learning under low effort (5 gm.) and the other under high effort (32 gm.). In the extinction period the acquisition conditions were fractionated into a total of four extinction conditions, and in the test for spontaneous recovery, the four extinction conditions were fractionated into a total of eight spontaneous recovery groups.

RESULTS

Original learning.—The number of trials to a latency criterion was used to compare the learning properties of the high- and low-effort groups. The learning criterion was reached when the total latency of *S*'s responses over five consecutive trials was less than 5 sec. The learning scores included the five criterion trials. The mean scores for the low-effort and high-effort groups were 22.90 and 27.63 trials, respectively ($t = 3.47$, $df = 78$, $P < .01$). The learning curves for the two groups are shown in Fig. 2. Although the low-effort group reached

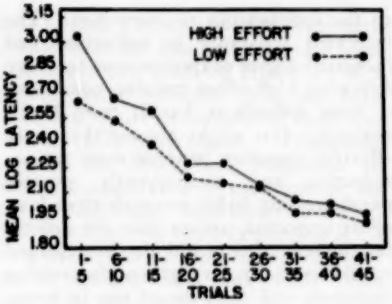

FIG. 2. Learning curves for the high- and low-effort groups.

the criterion sooner, the high-effort group showed a more rapid rate of learning, in terms of the steepness of the learning-curve slopes.

Experimental extinction.—A 2 × 2 analysis of variance was performed on the trials to the extinction criterion of two consecutive 30-sec. trials without a response. A Bartlett test indicated that homogeneous variance could be assumed. The means and *SD*s of these four groups are shown in Table 1 and the analysis of variance in Table 2. As can be seen, none of the between-groups sources of variance reach the conventional .05 level.

TABLE 1

MEANS AND *SD*s OF TRIALS TO EXTINCTION

Learning Level	Extinction Level	Mean Trials	*SD*
High	High	46.75	23.48
High	Low	56.50	32.46
Low	High	35.55	26.64
Low	Low	45.50	22.10

TABLE 2

ANALYSIS OF VARIANCE OF EXTINCTION SCORES

Source of Variance	*df*	*MS*	*F**
Effort levels during learning (L)	1	2464.20	3.34
Effort levels during extinction (E)	1	1940.45	2.63
L × E	1	.20	.0003
Within groups	76	737.54	
Total	79		

* The *F* ratio for $P = .05$ is 3.97.

TABLE 3

MEANS AND *SD*s OF TRIALS TO SECOND EXTINCTION (SPONTANEOUS RECOVERY)

Motivation Level			Mean Trials	*SD*
Learning	Extinction	Recovery		
Low	Low	Low	10.1	3.96
Low	Low	High	6.4	1.96
Low	High	Low	9.3	6.17
Low	High	High	5.9	2.70
High	Low	Low	17.8	9.46
High	Low	High	8.9	3.45
High	High	Low	9.9	4.87
High	High	High	5.6	1.85

Spontaneous recovery.—The trials to a second extinction were used as an inverse measure of spontaneous recovery. The means and *SD*s are shown in Table 3, and the analysis of variance in Table 4. The three major variables all reached the .05 level or better. Examination of the means for the various conditions reveals that spontaneous recovery is greater with high effort during learn-

TABLE 4

ANALYSIS OF VARIANCE OF SPONTANEOUS RECOVERY SCORES

Source of Variance	*df*	*MS*	*F**
Effort levels during learning (L)	1	137.81	5.13
Effort levels during extinction (E)	1	195.31	7.27
Effort levels during recovery (R)	1	515.11	19.18
L × E	1	101.25	3.77
L × R	1	46.51	1.73
E × R	1	30.01	1.12
L × E × R	1	44.39	1.65
Within groups	72	26.86	
Total	79		

* For $P = .05$, $F = 3.98$; for $P = .01$, $F = 7.01$.

ing, low effort during extinction, and low effort during the test for spontaneous recovery. All interaction terms fell short of the .05 level of significance, although L × E is significant at the .06 level.

One further analysis was performed to clarify the interpretation of the spontaneous recovery data. The relationship between the number of trials to extinction and the extent of spontaneous recovery was determined without regard to the effort variable. Eight product-moment correlations were run on the scores of the spontaneous recovery groups and the same groups' extinction values and the obtained coefficients were then tested for the significance of their departures from an hypothesized zero correlation. With 8 *df* for each of these coefficients, a correlation of .632 was required for significance. None of the correlations reached this level. The highest was − .410, and about half of the coefficients were negative and about half positive.

Discussion

Interest in the effort variable in learning has largely grown out of the theoretical importance given it in Hull's behavior system. Therefore, the data will first be examined in the light of Hull's specific constructs on effort.

Hull conceptualizes effort or work entirely from the point of view of its subtractive effects on performance. Both temporary inhibition (I_R) and permanent inhibition ($_sI_R$) subtract directly from the response potential ($_sE_R$). Habit ($_sH_R$) is presumably free of inhibitory influences. From this, the decrements in learning, in resistance to extinction, and in spontaneous recovery with increased effort at each of these stages is predictable. The decrement in spontaneous recovery following high-effort extinction represents the greater development of learned inhibition ($_sI_R$) during the extinction trials carrying over to the spontaneous recovery test. The increased resistance to extinction and increased degree of spontaneous recovery following high-effort training conditions is more difficult to handle from Hull's position. One might assume that more effortful responses produce more fatigue reduction and, consequently, greater panel-pressing habit strength than low-effort responses, except that the greater fatigue reduction would also be expected to differentially strengthen the resting responses and thus cancel out in terms of the performance potential remaining.

Recently, Maatsch, Adelman, and Denny (4), in discussing a similar finding in the Applezweig study (1), conclude that this effect may not be governed by inhibition processes at all, but rather is a transfer-of-training phenomenon. They argue that although the high-effort learning is adequate to low-effort extinction responses, the reverse is not true. Although some sort of effect of this variety may be involved in this study, there are some data that do not square with the hypothesis. Specifically, one would expect from this viewpoint that groups of *S*s which adapted to the low-effort condition during learning and then extinguished and recovered under the same level to be equal in performance to *S*s which adapted and maintained a high-effort level throughout. Thus, the means for the first and last groups in Table 3 should be the same. However, the difference between these means (4.5) is significant at the .01 level ($t = 3.08$). Additionally, if practice at the high-effort response was all that was involved, then one would expect the groups extinguishing under high effort to perform more efficiently during the spontaneous recovery test. The opposite result was obtained. These facts seem to favor a work inhibition hypothesis rather than a transfer hypothesis.

Returning to Hull, an explanation of this effect can be obtained by means of the stimulus intensity dynamism during learning construct (V_1). If it can be assumed that the feedback stimulation from the bar-pressing responses is analo-

gous to Hull's V, and that the intensity of this feedback was greater under high than low effort, then the prediction of a greater habit strength following high than low effort is predicted from the formula: ${}_{S}H_{R} = {}_{S}H_{R} \times V_{1}$. This would explain the heightened resistance to extinction and greater spontaneous recovery following high-effort learning. The fact that this line of reasoning leads to a prediction of superior performance during learning under high effort, which was not obtained, may be at least partially understood as the result of: (*a*) the greater likelihood of the low-effort response occurring "spontaneously" during the early part of learning, and (*b*) the concomitant development of greater inhibition in the high-effort group. If it is objected that the bar-pressing feedback stimulation did not precede the response and, therefore, was not strictly analogous to the stimulus intensity construct, it can be noted that the intertrial responding of the *S*s did precede the bar-pressing and that they were consistently observed to be close surrogates of the actual intensity of the bar-pressing response.

If the nearly significant interaction term in the spontaneous recovery data is confirmed in subsequent research, it will present a problem for current inhibition theory. This effect is associated with the presence or absence of a change in effort level from learning to extinction, greater recovery occurring with a change. Although there is little to support it in the rest of the data, some sort of discrimination learning factor seems to be operating here, i.e., discrimination between reinforcement and nonreinforcement conditions.

Additionally, response-produced cues from high- and low-effort responding may have taken on learned reinforcement or learned fatigue properties, depending on the state of affairs following the response, and since such cues were quite likely more numerous in the high-effort groups, they could have provided another source of differences between the experimental conditions.

In any event, it should be apparent that the rather complex set of relationships found in this study cannot be incorporated into any existing set of constructs dealing with inhibition alone; the problem is considerably more complicated.

Summary

A panel-pressing apparatus was employed to study the effects of varying the amount of effort expenditure required to respond upon the acquisition, extinction, and spontaneous recovery of an instrumental response. Two effort levels were used with rats as *S*s. Half of the *S*s were trained under high effort and half under low for the same number of trials. Then half the *S*s in each of the learning conditions were extinguished under high and half under low effort, creating four extinction conditions. Finally, half the *S*s in each of the extinction conditions were tested for spontaneous recovery under high and half under low effort, producing eight spontaneous recovery groups.

The results were: (*a*) The learning criterion was reached significantly sooner under low- than high-effort expenditure. (*b*) There was increased resistance to extinction following high-effort learning and reduced resistance to extinction with high effort during the extinction trials, though neither difference was statistically significant. (*c*) There was greater spontaneous recovery following high-effort learning, reduced recovery following high-effort extinction, and reduced recovery with high effort during the recovery test. (*d*) The interaction between effort during learning and effort during extinction approached statistical significance.

References

1. Applezweig, M. H. Response potential as a function of effort. *J. comp. physiol. Psychol.*, 1951, **44**, 225–235.
2. Horenstein, B. R. Performance of conditioned responses as a function of strength of hunger drive. *J. exp. Psychol.*, 1951, **44**, 210–224.
3. Hull, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.
4. Maatsch, J. L., Adelman, H. M., & Denny, M. R. Effort and resistance to extinction of the bar-pressing response. *J. comp. physiol. Psychol.*, 1954, **47**, 47–51.
5. Mowrer, O. H., & Jones, H. Extinction and behavior variability as functions of effortfulness of task. *J. exp. Psychol.*, 1943, **33**, 369–386.

(Received December 12, 1955)

Journal of Experimental Psychology
Vol. 53, No. 1, 1957

SELF-PACING PERFORMANCE AS A FUNCTION OF PERCEPTUAL LOAD[1]

R. CONRAD AND B. A. HILLE

Applied Psychology Research Unit, Cambridge, England

Several recent studies of sensorimotor skills have considered the case of multichannel displays in which the next signal might come from any one of a number of alternative sources. In some cases (**1, 5**) both paced and self-paced conditions were examined, but the perceptual load itself was not an experimental variable. This latter case was considered by Knowles, Garvey, and Newlin (**6**). The tasks used by all these authors were of the reaction-time variety in which *S* had to respond as soon as possible *after* the signal had occurred. Conrad (**2**), Wagner, Fitts, and Noble (**10**), and Mackworth and Mackworth (**7**), considered the load variable and used a task in which *S* was required to judge, not guess, in advance, at which of a varying number of sources the next signal would occur. But they only examined the paced condition. The present study fills in this gap. The task used was self-paced, and the number of signal sources was the experimental variable. The moment of arrival of signals could be predicted as a result of perceptual anticipation (**8, 9**).

When a series of randomly spaced signals was presented at a paced constant mean rate, and their moments of occurrence could be predicted, it was shown (**4**) that performance decreased linearly as a function of the number of signal sources. Using the same apparatus and eight signal sources, it was further shown (**3**) that when self-pacing was made available to *S*s their performance was better than when they were paced at a rate equal to their individually selected self-paced rate.

These results led to the supposition that if the number of signal sources (load) was varied and *S*s were self-paced, they would work slower with larger loads, the pace selected being such that a constant performance would be maintained irrespective of the load condition.

METHOD

The apparatus used was the Multi-dial Display, which was described in the papers cited (**3, 4**). Briefly, this is a panel of revolving pointers arranged in a 4×4 manner, with a corresponding response panel of spring-loaded switches. When pointers reach the 12 o'clock and 6 o'clock marks they stop, and this constitutes an error. Pressing the corresponding switch starts the pointer. If, however, the correct switch is held on while the pointer is on the mark, the pointer does not stop. This is a correct response. The task, under the condition of self-pacing, is to make as many correct responses and as few errors as possible in a given time.

The pointers all move at slightly different speeds. The effect of this is that considering all pointers together, the interval between successive signals has an approximately random distribution. Effectively there is no repeating pattern and the location of one signal cannot be predicted from the location of the previous one. It can, however, be determined at any moment by comparing relative speeds of the pointers and their distances to the next mark. Continuous judgment of this kind is forced upon *S* since he is instructed to press only one switch at a time.

All the pointers are driven through a gear link by a single motor. A potentiometer mounted beside the panel of switches controls the motor speed, and hence the pointer speeds. This

[1] This work was carried out with the financial assistance of the Medical Research Council, England. The *S*s were supplied by the Royal Navy.

TABLE 1

EFFECT OF LOAD ON SELF-PACED DRIVING RATE AND ON PERFORMANCE

Measure	Number of Signal Sources (Load)				
	8	10	12	14	16
Mean driving rate (signals/min.)	17.9	17.4	17.2	17.0	16.6
Correct responses/min.	14.6	13.4	12.6	11.1	9.9
Errors/min.	3.3	4.0	4.6	5.9	6.7
Integrated score	11.3	9.4	8.0	5.2	3.2

determines the rate at which signals occur. It must be emphasized that neither the motor speed nor the signal rate need be maintained at a constant value. The *S* is free to vary either of them continuously.

Summarizing, using the potentiometer as an accelerator, *S* drives the pointers with his left hand and so presents himself with signals. With his right hand he operates the switches one at a time, so as to keep all pointers revolving.

The 18 naval ratings who served as *Ss* practiced for 30 min. on each of eight days during which all load conditions were introduced. On Day 9 they worked for 5 min. each, in random order, with 8, 10, 12, 14, and 16 pointers.

Performance was recorded in two ways. First, correct responses and errors were integrated and the result continuously indicated on a large clock visible to *S*. Second, these events were separately recorded on a polygraph, their sum being the total number of signals presented. Clearly, if *S* drove too fast a large number of errors would occur. If he drove too slowly, few correct responses could be made. The optimum speed under the conditions described, was that which advanced the integrated score at maximum rate, i.e., the arithmetic difference between correct responses and errors.

RESULTS

Two measures are available. The self-paced driving rate in terms of signals per minute presented, and performance. The mean values of these as a function of load, are shown for 18 *Ss* in Table 1.

Superficially it would appear that the supposition which prompted the experiment, namely, that *S* would drive at a slower rate in order to maintain performance, is not confirmed. The differences between mean driving rates was tested by analysis of variance. The summary given in Table 2 indicates that these differences are too small to attain statistical significance, nor are the differences between *Ss* statistically significant. A third point to note about the data in Table 1 is that the slope of the best fitting straight line through the error scores is almost exactly that which would fit the same data from an earlier study (**4**) when *Ss* were paced.

The possibility remains that the failure to find differences between driving rates at different loads was due to an averaging effect in the trends for individuals. The departure of the individual trends from the horizontal was calculated, and when the mean of the departure was compared by *t* test with zero, it was found not to be significantly different.

It would appear then, that irrespective of load, all *Ss* tended to drive at a constant average rate, that this rate was not very different from one *S* to another, and that it was

TABLE 2

SUMMARY OF ANALYSIS OF VARIANCE OF MEAN DRIVING RATE

Source	*df*	Mean Square	*F*
Between subjects	17	38.99	.75
Between loads	4	4.30	.08
Remainder	68	52.02	

maintained in spite of worsening performance.

Discussion

Two separate questions arise from the results. First, by what means did *Ss* maintain a constant average driving rate? Second, why did *Ss* not drive at slower rates with increased load and maintain performance?

The one certain indication of performance that *Ss* had, was provided by the clock which integrated correct responses and errors. This would have shown objectively slower rates of progress with load increases. Yet this information appears largely to have been ignored.

At the same time, in considering what information was available which could provide a basis for maintaining a constant average driving rate, one is impressed by the formidable integrations which would appear to be necessary. For example, the sound of the main motor is clearly audible. If *Ss* thought that when the load was doubled they ought to drive the main motor at half the speed, and relied on changes in motor sound, the judgment would not be easy. But this is not the case. The *Ss* themselves continuously changed the motor speed as they adjusted to changing display conditions. To maintain an average of a half or two-thirds of the speed of a standard, these continuous changes, ranging from 0–1500 r.p.m., would need to be integrated. This is equally true of the position of the potentiometer speed control. Explanation along these lines does not seem to be feasible.

Alternatively it may be that *Ss* try to create an interval between successive signals which is equal within and between load conditions. Yet even if they were to achieve this, the integrating clock would show that it was unadaptive behavior, and this too would need to be accounted for. It would seem prudent therefore at this moment to leave these questions open, while reporting results which are relevant both to problems of pacing and to problems of multichannel displays.

Summary

Eighteen *Ss* were tested at a multichannel sensorimotor task under five conditions of load. The task required *S* to judge in advance which channel would provide the next signal and to respond accordingly, the interval between successive signals being near random. The task was self-paced, and *Ss* were instructed to drive at the rate which would give them the best score.

Analysis of variance showed that, irrespective of the load, *Ss* drove at the same rate in terms of signals per minute. But their performance was worse with larger loads, the trend being similar to that found previously under paced conditions. The effect occurred even though continuous knowledge of results was provided, while no direct information about driving rate was available. Differences between *Ss* were not significant.

References

1. Broadbent, D. E. Noise, paced performance and vigilance tasks. *Brit. J. Psychol.*, 1953, **44**, 295–303.
2. Conrad, R. Speed and load stress in a sensorimotor skill. *Brit. J. industr. Med.*, 1951, **8**, 1–7.
3. Conrad, R. The timing of signals in skill. *J. exp. Psychol.*, 1956, **51**, 365–370.
4. Conrad, R. Some effects on performance of changes in perceptual load. *J. exp. Psychol.*, 1955, **49**, 313–322.
5. Klemmer, E. T., & Muller, P. F., Jr. The rate of handling information. *Human Factors Operations Research Laboratories*, 1953. Memo. Rep. No. 34.
6. Knowles, W. B., Garvey, W. D., & Newlin, E. P. The effect of speed and load on display-control relationships. *J. exp. Psychol.*, 1953, **46**, 65–75.
7. Mackworth, J. F., & Mackworth, N. H. The overlapping of signals for decisions. *Amer. J. Psychol.*, **69**, 26–47.
8. Poulton, E. C. Perceptual anticipation in tracking with two-pointer and one-pointer displays. *Brit. J. Psychol.*, 1952, **43**, 222–229.
9. Poulton, E. C. The basis of perceptual anticipation in tracking. *Brit. J. Psychol.*, 1952, **43**, 295–302.
10. Wagner, R. C., Fitts, P. M., & Noble, M. E. Preliminary investigations of speed and load as dimensions of psychomotor tasks. *USAF, Personnel Train. Res. Cent., Res. Bull.*, 1954, No. 45.

(Received December 22, 1955)

Journal of Experimental Psychology
Vol. 53, No. 1, 1957

STIMULUS CONDITIONS AND RETROACTIVE INHIBITION

JOEL GREENSPOON
Florida State University

AND

REDGE RANYARD
Claremont, Calif.

Most of the research on retroactive inhibition has involved the manipulation of the materials to be learned. The investigation of the stimulus conditions under which learning occurred has received little attention. Two experiments which involved the problem of the stimulus conditions resulted in conflicting results. Nagge (4) failed to find a decrease in the amount of RI when interpolated learning occurred in a room different from that used during original learning and relearning. Bilodeau and Schlosberg (1), on the other hand, obtained a significant reduction in the amount of RI when the conditions in which interpolated learning occurred differed from the conditions in which original learning and relearning occurred.

The conflicting results are an adequate basis for repeating the experiment. In addition, Bilodeau and Schlosberg (1) noted that this problem has theoretical significance for the area of RI and learning theory. The interference theory of RI may be supported or refuted by the results of such investigations, since it places great emphasis on the interfering effects of the learning of one list of nonsense syllables on the learning of another list of nonsense syllables. However, if the stimulus conditions of OL and IL were differentiated, then the interfering effects of IL should be reduced. It was the purpose of this experiment to test this hypothesis.

METHOD

Apparatus.—Two sets of nonsense syllables of equal association values were selected from Glaze's (3) list of nonsense syllables. List 1 consisted of the following nonsense syllables: JIQ, GID, QUJ, KOJ, VUK, GEX, WOJ, XIQ, MEQ, and WUB. List 2 consisted of the following: JIC, XAJ, QIJ, CEF, MEF, VEH, GAX, YUJ, WEZ, and VAF. The association values of the nonsense syllables ranged from low to high. Each list contained an equal number of nonsense syllables of each association value. The order of the nonsense syllables in each list was randomly determined. An attempt was made to create high interlist similarity. Each list was typed in capital letters on separate sheets of paper and placed on separate drums. The same two lists were also printed by hand on cards, each syllable on a separate card. The cards were 8 in. by 6 in., and the letters were 5 in. high.

Two experimental conditions were used, the Card Room and the Drum Room. The Card Room was located in a basement room of the psychological laboratory of Pomona College. The room was about 10 by 8 by 8 ft. and was cluttered with old furniture and research equipment. The room was lighted by two 300-w. ceiling light bulbs. A table, 4 ft. high, was located about 6 ft. from a designated point in the room where *S* stood during the experiment. On the table was set a stand, 18 in. high, on which each card was placed. Behind *S*, connected by an open door, was the room where the departmental animal colony was maintained. The animal room was also brightly lighted during *S*'s presence. A door connected the main experimental room with a small room where a tape recorder was connected to a speaker. A tape contained the noise recorded from an FM tuner when off-station. A metronome was placed on a shelf alongside the table. A chair for *E* who recorded the responses of *S* was set off to the right of the table. An interval timer was also set on the chair.

The Drum Room was made as different as possible from the Card Room. It contained a small stand on which a memory drum was placed. The room was darkened except for a small white Christmas-tree light which was attached to the memory drum and lighted the small aperture through which one syllable at a time appeared. A soft chair was provided for *S*. A microphone, out of sight of *S*, was placed in the room and was attached to a speaker in

the adjoining room where *E* was located. A one-way screen enabled *E* to observe *S* during the experiment without being seen.

Design.—Although the two lists contained nonsense syllables of the same association values, the experiment was designed to control for any differences in the lists. The usual RI design was used, with original learning (OL), interpolated learning (IL), and relearning (RL), except for the manipulation of the stimulus conditions. Letting A stand for either the Card Room or Drum Room, and B for the other, the four basic conditions were AAA, AAB, ABB, and ABA. Sixteen groups of *S*s were used. For half of these groups A was the Card Room, and B the Drum Room. For the other eight groups it was reversed. Four of the groups within each eight had List 1 first and List 2 second; for the other four this was reversed.

Subjects.—The *S*s were 133 undergraduate students who volunteered. They were randomly assigned to 16 groups so that there were 8 *S*s in each of 11 groups and 9 *S*s in each of 5 groups.

Procedure.—The procedure was essentially the same for all groups. No preliminary practice was given. The *S* was given instructions to learn a list of 10 nonsense syllables and then proceeded immediately to learn the original list to a criterion of two consecutive errorless trials, after which he was given a 3-min. rest period. During the rest period, *S* was engaged in conversation by one *E* while the other *E* prepared the materials for IL. The IL was to the same criterion as OL. After IL, *S* was given a 3-min. rest period. The *S* relearned the original list to two consecutive errorless trials.

The method of anticipation was used in all three conditions of learning, and *S* was required to spell aloud each nonsense syllable. In OL and IL, however, *S* was permitted to look at the entire list of nonsense syllables once before beginning to anticipate and spell each one. In RL *S* began anticipating and spelling aloud each nonsense syllable on the first trial.

Each nonsense syllable was presented for 3 sec. with 3 sec. between each nonsense syllable. The timing was automatic in the Drum Room and by a ticking metronome in the Card Room. There were 57 sec. between presentations of the list.

The difference in the procedure involved preparing *S* for learning the nonsense syllables presented on the cards or memory drum. The instructions to *S*s who learned in the Drum Room were given outside the room. Instructions to *S*s who learned in the Card Room were presented under the conditions which prevailed during the learning trials. That is, *S* stood at the point where he was to stand during the learning trials, the lights were on, and the recorded noise was played. A second difference concerned informing *S*s who were to learn in the Drum Room that an asterisk preceded the first syllable in the list and *S*s who learned in the Card Room that a blank card preceded the first syllable. Another difference involved *E*. The *E* was not present in the Drum Room but was present in the Card Room, where *E* presented the cards to *S* and was within the visual field of *S* during the entire experimental session. In addition, *E* who recorded the responses of *S* was within the visual field of *S* in the Card Room but was not present in the Drum Room. The time interval between presentations of the list was automatically timed by the length of the list in the Drum Room. In the Card Room an interval timer was used which made a loud ringing noise when the time had elapsed. In the Drum Room *S* spent the rest interval outside the room in another room; in the Card Room *S* spent the rest interval in the room where the experiment was being conducted.

Results

The number of trials required to learn the original list was used to determine if either the rooms or the lists significantly affected original learning performance. The means and *SD*s of the number of trials to learn the original lists in the two different rooms are presented in Table 1. A 2×2 analysis of variance (2) showed that the difference between the means of neither the lists nor rooms was significant at the .05 level (F for lists = 3.45, $df = 1$ and 129, $P > .05$; F for rooms = 2.51, $df = 1$ and 129, $P > .05$). Since neither the lists nor the room significantly affected performance on original learning, the 16 groups were combined into 4 groups

TABLE 1

Number of Trials to Learn Original List of Nonsense Syllables

List	Drum Room		Card Room	
	Mean	*SD*	Mean	*SD*
1	21.23	7.13	19.91	8.87
2	19.91	6.01	16.39	6.08

TABLE 2

NUMBER OF TRIALS TO RELEARN AND NUMBER OF NONSENSE SYLLABLES RECALLED CORRECTLY ON FIRST RELEARNING TRIAL

Group	Trials to Relearn		Syllables Recalled	
	Mean	*SD*	Mean	*SD*
AAA	6.74	2.55	3.47	2.09
AAB	10.18	4.55	1.79	1.73
ABA	4.24	2.38	7.12	2.02
ABB	8.28	4.28	3.91	2.44

for the additional analyses. The four groups of *S*s who had OL, IL, and RL in the Drum Room or in the Card Room, whether they learned List 1 or 2 initially, were combined into Group AAA. Similarly, Groups AAB, ABA, and ABB were formed.

The means and *SD*s of the number of trials to relearn the original list are presented in Table 2. A simple analysis of variance (**2**) showed the differences among groups to be significant ($F = 16.23$, $df = 3$ and 129, $P < .001$). The mean differences between groups were analyzed by the t test for unmatched groups (**2**). The results of this analysis are presented in Table 3.

The means and *SD*s of the number of syllables correctly anticipated on the first RL trial are presented in Table 2. A simple analysis of variance showed the differences among groups to be significant ($F = 32.49$, $df = 3$ and 129, $P < .001$). The mean differences between groups were analyzed by the t test for unmatched groups, and the results of this analysis are also presented in Table 3.

TABLE 3

t-TEST ANALYSES OF MEAN DIFFERENCES BETWEEN GROUPS ON NUMBER OF TRIALS TO RELEARN AND NUMBER OF SYLLABLES RECALLED ON FIRST RELEARNING TRIAL

Group	AAB		ABA		ABB	
	Trials	Recall	Trials	Recall	Trials	Recall
AAA	3.78*	3.50*	4.10*	7.20*	1.77	0.77
AAB			6.53*	11.34*	1.72	4.00*
ABA					4.70*	5.63*

* $P < .001$.

Both response measures show that the conditions most favorable to recall and relearning consist of IL under conditions different from the conditions of OL and RL but with OL and RL occurring under the same conditions. The poorest conditions for recall and relearning consisted of OL and IL under the same conditions, but relearning the original list under different conditions.

To determine the effect of the conditions of OL on the learning of the interpolated list, the data were grouped on the basis of the interpolated list being learned under conditions which were the same or different from the conditions of OL. The measure used was the difference in the number of trials required to learn the original and interpolated lists. Both groups required fewer trials to learn the interpolated list. The mean gain in trials was 4.64 for the group that learned both lists under the same conditions and 3.36 for the group that learned the two lists under different conditions ($t = 1.00$). The difference could have arisen by chance.

DISCUSSION

The effect of the stimulus contexts on RI is clearly shown by the results of this experiment. Within the limitations of the similarity of the experiments, the results agree with those obtained by Bilodeau and Schlosberg (**1**). The results also tend to support the previously mentioned hypothesis.

Though the stimulus contexts differed

for the two conditions of learning, the two contexts could not be made completely different since the two lists were related in that they contained nonsense syllables of comparable association values. In addition, it is reasonable to assume that, with a criterion of two consecutive errorless trials, not all of the stimuli would be associated to the responses of the particular list. It could be expected that some stimuli not associated with the responses would be present in the relearning situation regardless of the interpolated activity. Thus, it could not be expected that all *Ss* in the ABA Group would have perfect recall on the first relearning trial. However, it could be expected that the amount of RI of the ABA Group would be minimized because different responses could not be associated with the stimulus context of OL since most of the stimuli present in the OL context were not present during the IL context. In the AAA Group, however, many of the stimuli associated with the interpolated list responses would be present during relearning. This situation would tend to increase the amount of RI. This same argument would suggest that OL and IL under different stimulus contexts would produce less proactive inhibition than learning the two lists under the same stimulus contexts. The results do not support this position. The *Ss* who learned the original and interpolated lists under the same conditions would have the advantage of having learned the two lists under the same conditions. They had learned a method of learning nonsense syllables in OL which would transfer to IL. The *Ss* who learned the second list under different conditions from the original list were confronted with conditions which permitted little transfer with regard to learning how to learn nonsense syllables. The *Ss* were naïve and had never been confronted with nonsense syllables prior to the experiment. It is hypothesized that *Ss* given pretraining in how to learn nonsense syllables under the conditions used in the experiment would show less interference from having learned an original list if the interpolated list was learned under different conditions. This hypothesis was strengthened by the more frequent request for repeating the instructions of the interpolated list by *Ss* who were learning it under conditions which differed from the conditions of OL. This occurred although the instructions, except for the previously indicated differences, were the same for both conditions of interpolated learning.

It is interesting to note that the ABB Group performed as well as the AAA Group. In this situation there was little interference from having learned the interpolated list since the interpolated list was learned under conditions which had few stimuli which were common to OL. However, the relearning occurred in a condition which contained few stimuli which were common to the situation in which OL occurred. The AAB Group would be the group which was provided with the conditions least conducive to good performance. Many of the stimuli associated to the responses of the original list were present and ultimately associated to the responses of the interpolated list. The responses learned to the interpolated list would tend to interfere with the responses learned to the original list. When placed in the relearning condition which differed from OL, *Ss* would be able to make few correct responses. Some *Ss* in this group reported that the list in RL seemed to be a new and different one and they required as many trials to relearn the list as they had required to learn the list originally.

Summary

The effect of the manipulation of the stimulus conditions on the recall and relearning of an originally learned list of nonsense syllables was investigated. A total of 133 *Ss* was randomly assigned to 16 different groups which involved four patterns of stimulus conditions (AAA, AAB, ABA, ABB) and their reversals during

original learning, interpolated learning, and relearning and counterbalancing of the two lists. The results showed that the group that learned the interpolated list under conditions which differed from the conditions of original learning and relearning (SDS) was superior to the three other groups in terms of recall and relearning measures. The group that relearned under conditions which differed from the conditions of original and interpolated learning (SSD) was inferior to the three other groups on both measures. The two other groups did not differ significantly from each other. Learning the interpolated list was not significantly affected by the similarity or difference of the conditions of original and interpolated learning.

References

1. Bilodeau, I. McD., & Schlosberg, H. Similarity in stimulating conditions as a variable in retroactive inhibition. *J. exp. Psychol.*, 1951, **41**, 199–204.
2. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
3. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–269.
4. Nagge, J. W. An experimental test of the theory of associative interference. *J. exp. Psychol.*, 1935, **18**, 663–682.

(Received December 12, 1955)

Journal of Experimental Psychology
Vol. 53, No. 1, 1957

RETROACTIVE INHIBITION AS A FUNCTION OF THE DEGREE OF ORIGINAL AND INTERPOLATED LEARNING[1]

GEORGE E. BRIGGS[2]

Northwestern University

Amounts of practice devoted to the original and to the interpolated learning sessions are major variables in the retroactive inhibition (RI) paradigm. McGeoch (**4**) found RI to be inversely related to the degree of original learning (OL), while several studies have consistently shown that RI increases initially as a function of the degree of interpolated learning (IL), reaches a maximum for intermediate amounts of practice on IL and may remain at a constant level or decrease slightly at higher degrees of the IL variable (**5, 8, 10, 11**). In these studies when the degree of IL was varied, OL was held to a constant value, and vice versa.

The present study was designed to measure RI as a function of the degree of IL for several levels of practice on OL. Further, the relative strengths of the OL and IL response classes were determined just prior to the traditional measure of RI.

METHOD

Subjects and materials.—Eighty *Ss* served in this experiment as part of the requirements for classes in introductory psychology at the Northwestern University. Each *S* served a total of five days for approximately 45 min. per day. These *Ss* learned two lists of 10 paired adjectives, OL and IL, in the reproductive inhibition paradigm during each daily session, and both learning and recall occurred in the same session. The lists were presented on a modified Hull-type memory drum which was adjusted for a 2-sec. rate of exposure. There were three orders of presentation for each list, and an 8-sec. interval occurred between each trial.

Design and procedure.—On Day 1 *S* was given instructions pertinent to the learning of paired associates. He then learned two practice lists to a criterion of one perfect trial. One minute after IL each *S* received a modified free recall (MFR) test followed by five relearning trials on the original list. During MFR *S* was presented with the 10 stimulus words singly, and he was requested to give the first of the two responses to come to mind. If neither response was available for recall, *S* was urged to give *any* response word. In this way *E* was able to measure response dominance without penalizing *S* for "incorrect" responses.

On each of the subsequent four days, *S* received certain numbers of trials on the pairs of experimental lists. There were four degrees of learning allotted for OL (2, 5, 10, and 20 trials) and five degrees of learning for IL (0, 2, 5, 10, and 20 trials). The design of the experiment utilized five independent groups of 16 *Ss* each. Trials IL was the group variable. The *Ss* randomly assigned to each group also were assigned in a nonsystematic manner to sequence of treatments, Trials OL, within groups. A standard 4×4 greco-latin square was selected (**2**) and a random square generated therefrom. The Greek letters designated Trials OL while the Latin letters denoted lists of adjective pairs.

In this design a given *S* received the same number of trials on IL during each daily session but received all four values of the OL variable over the four experimental sessions. Use of the greco-latin square made the four pairs of experimental lists orthogonal to both experimental variables, Trials OL and Trials IL. The same square was used for each of the five IL groups.

Table 1 shows the daily treatments for the five IL groups. As indicated, a constant period of time elapsed between the termination of OL and the retention and relearning trials. Since the *Ss* in Group 5 received the greatest number of experimental trials, it was necessary to equate practice effects over the experimental period.

[1] This research was completed while the author was a postdoctoral fellow of the National Science Foundation. The author wishes to express his gratitude to Prof. B. J. Underwood for his advice and encouragement during the planning and conduct of this study.

[2] Now at the Laboratory of Aviation Psychology of the Ohio State University.

TABLE 1

DAILY TREATMENTS FOR THE FIVE IL GROUPS

Group	OL	Rest	IL	Rest	MFR	Relearning	Practice
1	2, 5, 10, or 20 trials	17½ min. rest			1 trial	5 trials	20 trials
2	2, 5, 10, or 20 trials	30 sec.	2 trials	15½ min.	1 trial	5 trials	18 trials
3	2, 5, 10, or 20 trials	30 sec.	5 trials	13 min.	1 trial	5 trials	15 trials
4	2, 5, 10, or 20 trials	30 sec.	10 trials	9 min.	1 trial	5 trials	10 trials
5	2, 5, 10, or 20 trials	30 sec.	20 trials	1 min.	1 trial	5 trials	0 trials

Column 7 of Table 1 shows how this was accomplished. The practice lists were presented to Groups 1 through 4 following the relearning trials. A given adjective appeared in only one pair of experimental or practice lists, and a total of four pairs of experimental and four pairs of practice lists were used.

RESULTS

Table 2 presents the average number of correct anticipations per *S*, made during OL for each of the four degrees of the OL variable; these data are categorized according to IL groupings. The data defining the five means for each degree of OL were subjected to simple between-within analyses of variance. All four *F* ratios were at or less than unity thereby indicating no differential amounts of learning in OL for each of the OL groupings.

Table 3 presents the mean number of OL responses correctly anticipated during the first relearning trial (traditional recall) for each degree of practice on OL and IL. The data defining these means were subjected to an analysis of variance following transformation to arc sin (per cent recalled)$^{\frac{1}{2}}$ and the results are summarized in Table 4.

The significance of Sessions in Table 4 indicates that a practice effect was in evidence for the recall of OL responses in this experiment. The significance of Lists does not indicate a differential effect on the two experimental variables Trials OL and Trials IL for, as mentioned above, these three variables were orthogonal, one to another, in this design. The sources Trials OL, Trials IL, and their interaction are of special interest: all are statistically significant when tested against the appropriate error estimates. The Duncan Multiple Comparisons Test (1) was used to determine the pattern of significance between the 20 means involved in the Trials OL × Trials IL inter-

TABLE 2

AVERAGE NUMBER OF TOTAL CORRECT ANTICIPATIONS, PER *S*, DURING OL

Trials OL	Trials IL				
	0	2	5	10	20
2	4.75	5.44	5.62	4.69	2.94
5	17.94	15.69	21.69	16.56	17.68
10	51.12	49.25	54.50	46.50	48.31
20	143.25	131.50	137.38	140.31	144.50

TABLE 3

MEAN NUMBER OF OL RESPONSES ON FIRST RELEARNING TRIAL

Trials OL	Trials IL				
	0	2	5	10	20
2	3.25	1.94	0.94	0.94	0.94
5	4.94	2.38	2.12	2.06	1.62
10	7.19	5.44	4.19	3.12	3.06
20	9.44	7.88	6.44	4.75	4.38

Note: $\sigma_m = .41$.

TABLE 4

ANALYSIS OF VARIANCE OF TRADITIONAL RECALL DATA

Source	*df*	MS	*F*
Sequence	3	1088.35	2.29
IL	4	8975.33	18.85*
IL × Seq.	12	427.05	—
Ss/Seq./IL	60	476.05	
OL	3	22389.40	144.94*
IL × OL	12	345.99	2.24*
Sessions	3	754.94	4.88*
IL × Ses.	12	127.84	—
Lists	3	1668.36	10.80*
IL × Lists	12	122.12	—
Square Uniqueness	6	151.82	—
IL × Sq. Uniq.	24	114.81	—
Ses. × *Ss*/IL	165	154.47	

* Significant at or beyond $P < .05$.

action, and the results of this test are shown in Table 3. Those means within brackets are not significantly different, one from another, whereas those not so bracketed do differ at the $P < .05$ level.

In general, an increase in Trials OL is accompanied by an increase in the number of means differing significantly. This trend is illustrated by Fig. 1 which presents the mean number of OL responses correctly anticipated during the first, third, and fifth relearning trials. Immediate concern is with the data from the first relearning trial. As noted in previous studies (**5, 8, 10, 11**), recall of the OL response class is greatest for the control group (IL-0) with orderly decreases in recall as Trials IL increases. In agreement with the results of the Duncan test, traditional recall as a function of Trials IL shows an extension of slope as Trials OL increases; and finally, the functions show a progressive increase in over-all level as Trials OL increases.

The functions for relearning trials three and five in Fig. 1 demonstrate the transitory nature of RI: recall

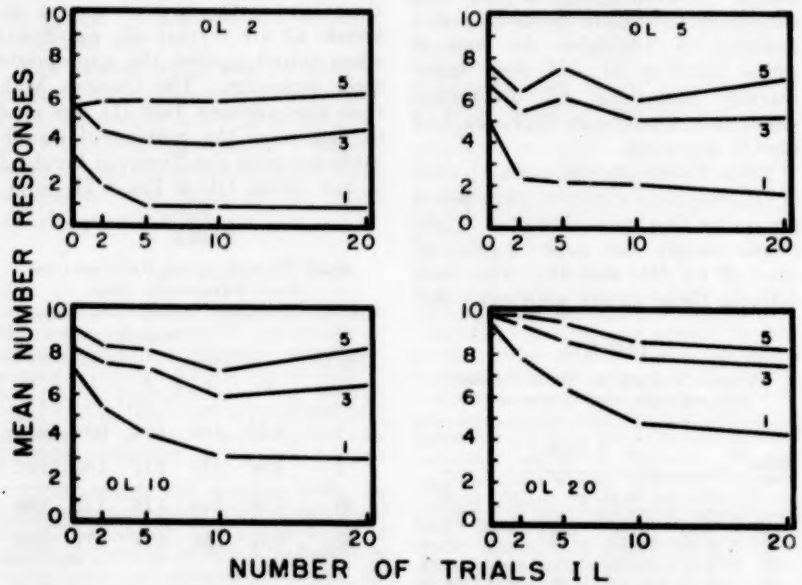

FIG. 1. Traditional recall and relearning functions.

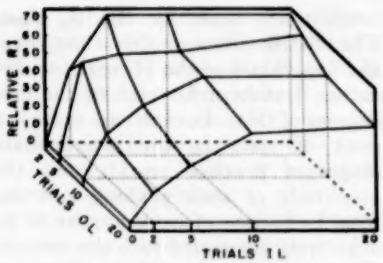

FIG. 2. Relative retroactive inhibition as a function of the degree of original and interpolated learning.

differences between the values of the IL variable decrease markedly from the first through the fifth relearning trials. Unlike the results of Thune and Underwood (**10**), RI here dissipates more rapidly under the IL-20 condition than under IL-10 for treatments OL-2, -5, and -10. The three relearning functions for OL-20 in Fig. 1 show almost equal rates of dissipation for conditions IL-10 and -20. These observations do agree with those of Melton and Irwin (**8**) where dissipation of RI was more rapid at both an IL-20 and an IL-40 condition than at a more intermediate stage of practice, IL-10. Therefore the results here confirm a conclusion reached by Melton and Irwin: ". . . RI disappears so much more rapidly with overlearned material than it does with moderately learned interpolated material" (**8**, p. 187). As indicated above, however, such differential dissipation appears only for low and intermediate degrees of practice on OL.

The data of Table 3 were used to determine relative RI, [Rest—Work/Rest] × 100, and these results are shown in Fig. 2. As with previous data, RI increases rapidly for initial increments in Trials IL and then levels at the intermediate stages of practice on IL (**5, 8, 10, 11**). As would be

TABLE 5

NUMBER OF INTRUSIONS ON FIRST RELEARNING TRIAL CLASSIFIED AS TO RESPONSE CATEGORY

Trials OL	R Cat.	Trials IL				
		0	2	5	10	20
2	R_2	0	12	26	12	1
	R_0	3	3	0	4	0
	R_3	8	14	5	10	1
5	R_2	0	11	23	14	0
	R_0	2	7	3	1	2
	R_3	5	7	8	4	6
10	R_2	0	5	10	10	4
	R_0	0	2	2	0	2
	R_3	8	9	13	7	3
20	R_2	0	1	7	9	1
	R_0	0	2	0	0	0
	R_3	7	8	12	5	7

expected from the McGeoch data (**4**), the over-all level of RI and the initial slopes of the RI functions decrease with increases in Trials OL. At the point OL-5, IL-2, the function appears to be elevated above that to be expected from the levels of the several neighboring points. It would appear

TABLE 6

AVERAGE NUMBER OF RESPONSES FROM MODIFIED FREE RECALL CLASSIFIED AS TO RESPONSE CATEGORY

Trials OL	R Cat.	Trials IL				
		0	2	5	10	20
2	R_2	—	1.62	3.81	6.31	8.31
	R_1	3.44	2.12	1.00	0.62	0.19
	R_{0-3}	6.56	6.25	5.19	3.06	1.50
5	R_2	—	1.56	3.50	5.31	7.69
	R_1	5.06	2.81	2.25	1.50	0.75
	R_{0-3}	4.94	5.62	4.25	3.19	1.56
10	R_2	—	1.00	2.50	6.50	7.56
	R_1	7.00	5.88	3.56	2.06	1.00
	R_{0-3}	3.00	3.12	3.94	1.44	1.44
20	R_2	—	0.56	2.06	4.88	6.44
	R_1	9.31	7.31	5.94	4.00	2.25
	R_{0-3}	0.69	2.12	2.00	1.12	1.31

that sampling error could account for this result.

Under the traditional method of measuring RI, the index to competition between response classes is furnished by the overt intrusions made by *S* during the first relearning trial. The frequency of these intrusions, classified as to response class, is presented in Table 5. A response was classified R_2 if it was a correct response in IL, and therefore the R_2 category represents direct intrusions of an IL response for an OL response. An R_0 response was one which did not appear in either OL or IL in either a stimulus or a response capacity, and therefore these responses were original with *S*. An R_3 response did appear in OL or in IL but was an incorrect response, in either list, when so anticipated during the recall trial.

In Table 5 the data of primary interest are those for the R_2 class. The distributions of this class, over the five values of the IL variable, are rather consistent for each of the four degrees of OL. In each case there is a peak of error in the intermediate stages of practice on IL, and the amplitude of such peaking and the total frequency of error appear to be negatively correlated with the amount of practice devoted to OL. One cannot emphasize the shapes of these distributions because of the low frequencies involved. However, there is consistency both with previous data (**8, 10**) and within the present data.

Attention is now directed to the responses on the modified free recall (MFR) test which preceded traditional recall and relearning in each experimental session. These data were categorized as to response class and then averaged over the 16 *S*s of

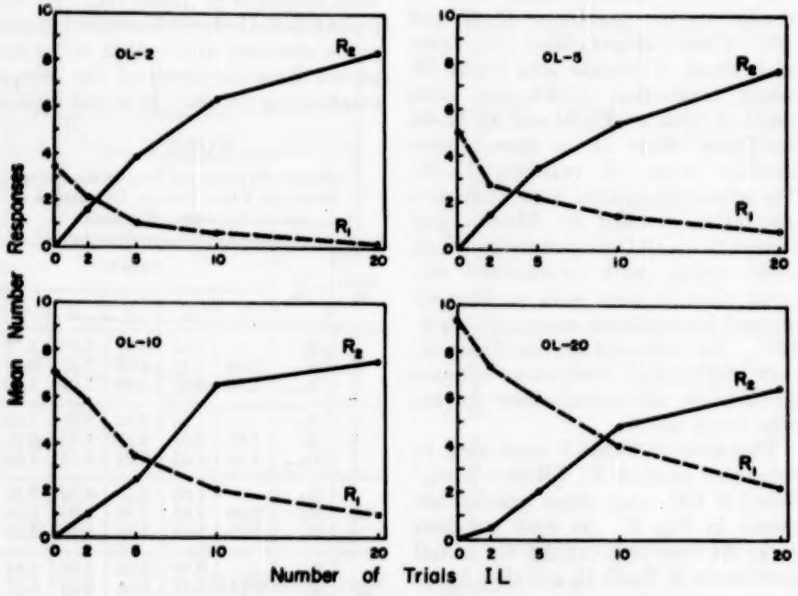

FIG. 3. Relative response strengths of the OL and IL response systems as measured by modified free recall.

each OL-IL combination. These averages are presented in Table 6 where response classes R_0, R_2, and R_3 are as defined earlier while R_1 refers to those responses correctly recalled during MFR from the OL lists. Response classes R_0 and R_3 have been combined in this table.

The data for response classes R_1 and R_2 are of primary interest here. For each OL grouping the frequency of R_1 responses decreases regularly as a function of Trials IL while R_2 responses increase. This is shown graphically in Fig. 3. For OL-2, R_2 clearly emerges as the dominant response system at IL-5 while for OL-5, R_2 has barely achieved dominance at IL-5 and it is not until IL-10 that such superiority is clearly apparent. Further, for OL-10 clear dominance of R_2 over R_1 does not occur until IL-10, and for OL-20 unequivocal superiority of R_2 does not appear until IL-20. The relevance of these results will be discussed in the following section.

Discussion

Previous research on RI as a function of the degree of IL not only has defined this relationship empirically, but in addition each study has contributed to a discussion of a theoretical basis for this RI. McGeoch was the first to name competition between OL and IL responses at the time of recall as the basis for RI (6), and subsequent reports have favored this factor in one form or another. For example, Melton proposed a two-factor theory of RI with competition of response during relearning as one factor and unlearning of OL responses during IL as the second (8, 9). Thune and Underwood preferred to account for RI by changes in the covert-partial-overt intrusions ratio (10), a view which places major emphasis on competition of response without denying the influence of an unlearning factor. Finally, it has been pointed out (6) that unlearning can be interpreted as an end product of competition, and therefore this factor emphasizes that such competition occurs during IL as well as during relearning.

The frequency of R_2 intrusions has been accepted as the index to competition of response at the time of recall. Since the MFR data of Fig. 3 describe certain of the conditions which make for intrusions at recall, a consideration of these data is warranted. Thus a comparison of the R_2 intrusion data of Table 5 and the relative positions of the R_1 and R_2 functions of Fig. 3 shows that a maximum number of such intrusions occurred at or near the point on the IL variable after which the R_2 function first emerged as dominant over the R_1 class of response. This is comparable to the results reported by Melton from a study in which the greatest frequency of R_2 intrusions during relearning occurred when OL and IL were approximately equally well learned (7). Further, in the present study, the initial increases in the frequencies of R_2 intrusions are similar to the initial slopes of R_2 as these latter were measured by MFR.

These observations suggest a dependence of the frequency of R_2 intrusions, as measured by traditional recall, on the strength of the R_2 system relative to the R_1 system of response, as measured by MFR, for the low and intermediate stages of practice on IL. However, the R_2 class continues to gain in strength over the R_1 class for higher degrees of learning in IL, and therefore one may assume an additional factor to account for the decrease in R_2 intrusions at those higher levels of IL. It would appear that the concept of differentiation, as proposed by Gibson (3), would be appropriate here. The differentiation hypothesis indicates that intrusion frequency (error) should be high whenever the OL and IL response systems, R_1 and R_2 respectively, are of near equal strength as under this condition, the two systems have low discriminability.

Three logical deductions follow from this hypothesis and may be tested by the

present data. First, maximum intrusion frequency should occur at or near the point where the two response systems are of nearly the same strength regardless of the absolute level of practice. Second, for low levels of the OL variable, increase in the IL variable above the point of equal strength should be accompanied by a decrease in intrusion frequency due to the increase in discriminability between R_1 and R_2 which results from differential levels of practice on the two lists. Finally, the total intrusion frequency should decrease for intermediate and high levels of practice on OL because again, discrimination is increased by the differential practice on the R_1 and R_2 systems of response.

These deductions are confirmed by the intrusion data of Table 5 and the relative dominance functions of Fig. 3. For both OL-2 and OL-5, increases in intrusion frequency follow increases in the strength of the R_2 response class, Fig. 3, to the point where R_2 has emerged as the dominant system, IL-5 in both cases. Further increase in R_2 is then accompanied by a reduction in overt errors, and this would be expected from the second deduction, above.

For OL-10 and OL-20, the R_1 class has achieved considerable strength, as compared to OL-2 and OL-5, prior to the IL session, and therefore the two response classes R_1 and R_2 should be more clearly discriminable at the low levels of the IL variable. Thus the third deduction is confirmed by the lower frequency of intrusions at the points IL-2 and IL-5 for conditions OL-10 and OL-20. Finally, it is seen that the peak of error has shifted from IL-5 for conditions OL-2 and OL-5 to the point IL-10 for conditions OL-10 and OL-20, thereby confirming the first deduction.

Thus the relative strengths of two major systems of response are related to the frequency of overt errors at the time of recall in a manner predictable from the Gibson hypothesis. It would appear, then, that this factor of differentiation should be considered in theoretical analyses of RI as a function of amounts of practice devoted to original and to interpolated learning.

Summary

Eighty *Ss* served for five consecutive days in a study of the degree of OL and the degree of IL in the reproductive inhibition paradigm. There were four degrees of learning for OL (2, 5, 10, and 20 trials) and five degrees for IL (0, 2, 5, 10, and 20 trials). The OL and IL lists contained 10 paired adjectives, and two retention measures were used: modified free recall followed by traditional recall and relearning. There were five relearning trials for the original list. The basic design was a 4×4 greco-latin square with Trials OL as the Greek letters and with lists as the Latin letters. There were five groups of 16 *Ss* each, and Trials IL was the group variable. The *Ss* were randomly assigned to groups and to sequence of treatments (Trials OL) within groups.

The results are as follows:

1. For each degree of OL, the five IL groups did not differ significantly in terms of the learning of the OL lists.

2. Frequency of OL recall decreased as a function of the degree of IL. As degree of OL increased, these recall functions became more extensive in slope and the over-all level of recall increased.

3. Differential RI as a function of Trials IL decreased markedly over the five relearning trials. Such decrease was more rapid for IL-20 than for IL-10 with treatments OL-2, -5, and -10.

4. The data from modified free recall indicated that the frequency of direct, interlist intrusions was closely related to the relative strengths of the OL and IL response systems. This was shown to be predictable from the Gibson hypothesis of differentiation.

References

1. Duncan, D. B. A significance test for differences between ranked treatments in an analysis of variance. *Virginia J. Science*, 1951, **2**, 171–189.
2. Fisher, R. A., & Yates, F. *Statistical tables*. New York: Hafner, 1949, 3rd Ed.
3. Gibson, E. J. Retroactive inhibition as a function of the degree of generalization between tasks. *J. exp. Psychol.*, 1941, **28**, 93–115.
4. McGeoch, J. A. The influence of degree of learning upon retroactive inhibition. *Amer. J. Psychol.*, 1929, **41**, 252–262.
5. McGeoch, J. A. The influence of degree of interpolated learning upon retroactive

inhibition. *Amer. J. Psychol.*, 1932, **44**, 695–708.

6. McGeoch, J. A. *The psychology of human learning.* New York: Longmans, Green, 1945.
7. Melton, A. W. Overt interlist intrusions and retroactive inhibition as a function of the ratio of the degrees of learning of original and interpolated verbal habits. *Psychol. Bull.*, 1941, **38**, 575. (Abstract)
8. Melton, A. W., & Irwin, J. McQ. The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *Amer. J. Psychol.*, 1940, **53**, 173–203.
9. Melton, A. W., & Von Lackum, W. J. Retroactive and proactive inhibition in retention: evidence for a two-factor theory of retroactive inhibition. *Amer. J. Psychol.*, 1941, **54**, 157–173.
10. Thune, L. E., & Underwood, B. J. Retroactive inhibition as a function of degree of interpolated learning. *J. exp. Psychol.*, 1943, **32**, 185–200.
11. Underwood, B. J. The effect of successive interpolations on retroactive and proactive inhibition. *Psychol. Monogr.*, 1945, **59**, No. 3 (Whole No. 273).

(Received December 19, 1955)

Journal of Experimental Psychology
Vol. 53, No. 1, 1957

THE RELATION OF DECISION-TIME TO STIMULUS SIMILARITY [1]

WILLIAM N. DEMBER

University of Michigan

The primary purpose of the present paper is to demonstrate that the experimental psychologist, given the proper methodology, can meaningfully investigate the unreplicated behavior of a single individual. In the experiments reported here *decision-time* is the dependent variable; it is shown that the rank order of the members of a set of decision-times, each representing a single measurement from one *S*, can be successfully predicted from the similarity of the alternatives about which the decisions are made. That such a delicate variable as decision-time is amenable to prediction in this manner is of especial interest to behavior theory, as is for example implied in a recent article by Spence (**10**).

The relation between decision-time and similarity was studied as early as 1902 by Cattell (**2**), who suggested that decision-time could provide data otherwise unobtainable via conventional "scaling" methods. He argued that *S* could not with any certainty judge whether Stimuli A and B were more or less similar than Stimuli C and D; *S* could, however, discriminate between A and B and between C and D. The relative times for these two decisions should indicate the relative distances between the two pairs of stimuli: the more similar pair would require the longer decision-time.

The validity of Cattell's hypothesis has been the subject of a great deal of experimentation, the purpose of which has been, in general, to demonstrate that decision-time is predictable from similarity as measured independently. With that established, it would then be possible to use decision-time in the absence of other information about similarity.

All of the above experimentation required a method for the measurement of similarity. The value of these experiments is limited, then, by certain characteristics of the particular method used. Ideally, the method would be based on the consistent responses of a single *S*, thus requiring neither averaging over *S*s nor over replications on one *S*. It would further be desirable that the responses from which the similarity assessment is made are the very responses yielding the decision-time data. With the recent development of measurement theory by Coombs (**3, 4, 5, 6**) methods have come into existence which meet all these requirements.

The present paper reports two experiments in which Cattell's hypothesis has been tested through the utilization of one of these new methods, in particular, one of the variants of the Method of Cartwheels. This particular technique (The Method of Cartwheels, Disjoint Rims) as a method of *collecting* data may be thought of as a variation of the classical Method of Paired Comparisons. The *S* is presented with four stimuli, arranged as two pairs (rims). His task is to choose the rim

[1] This paper is based on a doctoral dissertation submitted to the graduate school of the University of Michigan. The author is indebted to E. L. Walker and C. H. Coombs for their suggestions and criticisms. The research was conducted while the author held a predoctoral fellowship of the Social Science Research Council.

whose members are the more alike. In effect, he is comparing two *distances* and selecting the smaller. This, of course, is the very method considered unusable by Cattell.

The responses obtained in this manner may be analyzed by a special case of the Unfolding Technique (**3, 6**) to determine (*a*) whether the stimuli fall on a single dimension, (*b*) if so, how they are ordered on that dimension, and (*c*) what the relative distances are between pairs of stimuli.

Given the above information about the stimuli it is then possible to assign to each cartwheel a value, symbolized in this paper by *d*, representing the distance between the two rims of the cartwheel. These *d* values may themselves be ordered. It is hypothesized that these *d* values should be negatively correlated with the decision-times associated with the original cartwheels judgments.[2]

Method

Apparatus.—Two experiments are reported here. The stimuli of Exp. 1 were five patches of Munsell gray paper, referred to henceforth as A, B, C, D, E. Their luminance values were respectively .914, .522, .261, .207, and .109 ft.-lamberts in the experimental apparatus. The particular method of cartwheels which was used called for the stimuli to be presented four at a time, arranged in the configuration of a left- and right-hand pair. To accomplish this a special "template" was constructed, which served as the exposure field of a Gerbrands tachistoscope. It consisted of a white cardboard faced with white paper, into which four 1-in. square holes had been cut symmetrically about the center. The two members of each pair were separated by ½ in. The left- and right-hand pairs were separated by a distance of 1 in. Through a simple slot arrangement in the back of the card *E* could rapidly insert and remove the patches of gray paper from trial to trial.

For Exp. 2, a very similar template was prepared, differing only in the separation between left and right and top and bottom. The distance between left-right pairs was 2 in., and the top-bottom separation was 1 in. The stimuli (the words *always*, *often*, *rarely*, *seldom*, and *never*) were printed in India ink on white paper.

The *S*'s task was to report as soon as he had decided in which pair, left or right, the members were more similar—in Exp. 1 with respect to brightness, in Exp. 2 with respect to meaning. The *S* indicated his choice by depressing either of two microswitches. The closing of either switch served to stop a chronoscope which was automatically started with the presentation of the stimuli. The time interval between the presentation of the stimuli and the *S*'s response could be read with precision to .01 sec. The *S*'s choice, left or right, was indicated to *E* by two lights, each of which was activated by one of *S*'s microswitches.

Subjects.—The *S*s were four University of Michigan students, three men (D. C., V. C., M. H.) and one woman (P. K.). All had previously taken part in at least one psychological experiment. While all four were used in Exp. 1, only P. K. served in Exp. 2.

Procedure.—Several steps preliminary to the test session were taken in an attempt to obtain reliable choices and decision-times. These included giving *S* a brief explanation of Coombs' measurement theory, as well as at least two days' practice on a task similar to that of the experiment. Further, all but P. K. were given a day of practice on the experimental task itself. Finally *S* was informed that while there were no right or wrong answers, consistency was desirable; to insure consistency he should not think of anything else but the task during each trial, and report his decisions immediately.

The 40 cartwheels used in the test session were presented in random order, at the rate of about 1 per min. Before each trial *E* signalled verbally that a presentation was imminent, and then delayed presentation until about ½ sec. after *S* had signalled that he was ready.

Results

Experiment 1.—In Table 1 are the 40 cartwheels, the responses made to them, and the associated decision-times. From the responses made by each *S* it is possible by means of the Unfolding Technique to determine for him the order of the stimuli and of the distances between them. From this information a partial order of the *d*s may be derived for each *S*.

The hypothesis to be tested calls

[2] For a more complete description of the preceding technique see Coombs (**3, 4, 5, 6**) and Dember (**7**).

for a negative correlation between decision-time and *d*. To accomplish this test one further step in the data analysis was necessary, resulting from the fact that the set of *d*s is only partially ordered since some pairs of *d*s are incomparable. The following procedure was used. The partially ordered set (poset) of *d*s was examined. It was determined what the longest, unbroken "path" was through the poset. All paths which were equally as long (in terms of the number of elements, or *d*s) were recorded. These different paths were not necessarily independent; in fact, for the posets of this paper, they all had elements in common.

TABLE 1
CHOICES AND DECISION-TIMES (SECONDS) FOR EACH CARTWHEEL OF EXP. 1

Cartwheel*		Choice	Subjects			
L	R		D. C.	V. C.	M. H.	P. K.
AC	BC	R	1.62	4.20	2.28	2.61
AD	BD	R	1.91	1.36	2.90	3.54
AE	BE	R	1.70	1.33	1.68	7.09
AB	AC	L	1.38	3.11	2.21	2.62
BD	CD	R	1.32	1.62	1.36	3.35
BE	CE	R	1.63	2.38	1.61	2.31
AC	AD	L	2.75	1.94	9.93	5.52
BC	BD	L	2.41	2.59	6.38	6.98
CE	DE	R	3.33	6.04	4.41	11.53
AD	AE	L	1.51	3.10	2.87	2.25
BD	BE	L	2.31	3.98	2.43	3.31
CD	CE	L	1.39	2.83	1.45	3.09
AD	CD	R	1.05	1.72	1.76	3.01
AE	CE	R	1.31	3.22	1.24	2.84
AB	AD	L	1.52	1.52	2.53	4.19
BE	DE	R	1.30	4.03	4.89	2.73
AC	AE	L	1.45	2.00	2.82	3.13
BC	BE	L	1.62	2.40	1.94	3.88
AE	DE	R	1.07	1.08	1.15	2.39
AB	AE	L	1.50	1.84	1.49	2.01
AD	BC	R	1.81	2.55	1.93	2.07
AE	BC	R	1.69	2.25	1.45	2.93
AE	BD	R	1.15	1.31	2.10	2.50
AE	CD	R	1.20	1.65	1.16	2.20
BE	CD	R	1.28	2.45	2.08	2.75
AB	BC	L	3.37	4.83	*16.93*	12.81
AB	BD	L	2.34	2.57	3.54	5.44
AB	BE	L	1.66	1.78	2.29	3.40
AB	CD	R	2.27	3.45	5.53	4.19
AB	CE	L	3.77	6.20	*29.17*	11.59
AB	DE	R	2.54	*7.82*	21.91	7.33
AC	CD	R	1.30	1.72	1.19	3.14
AC	CE	R	1.52	1.66	9.36	6.58
AC	DE	R	1.74	2.61	1.30	1.90
AD	DE	R	1.40	1.26	0.94	4.42
BC	CD	R	1.91	3.48	2.25	3.33
BC	CE	L	3.77	3.20	*31.56*	6.75
BC	DE	R	3.65	5.56	16.94	7.77
BD	DE	R	2.70	3.36	6.05	6.65
CD	DE	L	2.13	3.70	3.25	4.75

* The rims of the cartwheels are labeled L (left) and R (right); the first of each pair of letters indicates the top member of the rim. The choices which differed from those given are indicated by italicizing the associated decision-time.

A rank correlation (tau) was then computed between each of these simple orders of *d*s and the orders of the appropriate decision-times. The taus computed between *d* and decision-time for each of the simply-ordered subsets of *d*s form a set. It has been decided that the most efficient treatment of this set of taus is to select from it the highest, lowest, and median values. These are presented in Table 2, along with the probability of obtaining each tau by chance. It can be seen from this table that the hypothesis has been sustained at a high level of confidence for each of the four *S*s.

For the stimuli of this experiment it was possible to obtain *physical* as well as psychological "values." It was considered of some interest to examine the relation between the perceived and the physical distances between stimuli and also the relation between decision-time and physical distance between rims. Since psychophysical functions are often best described logarithmically, it was decided to examine the above relations using log ft.-lamberts as the basis of the physical assessment.

The log ft.-lamberts assessment of distance between stimuli and the cartwheels assessment were highly correlated for all *S*s. For D. C. tau equaled .91, for V. C., 1.00, for M. H., .73, and for P. K., .91. The respective *P* values for the taus are: .00002, .0000003, .002, and .00002.

TABLE 2

RELATION BETWEEN DECISION-TIME AND *d*, EXP. 1

	[redacted] Subjects							
	D. C.		V. C.		M. H.		P. K.	
	tau	*P*	tau	*P*	tau	*P*	tau	*P*
Low	−.56	.001	−.24	.15	−.68	.0003	−.59	.0006
High	−.63	.0003	−.76	.0004	−.75	.00006	−.68	.00009
Median	−.56	.001	−.64	.002	−.73	.00009	−.62	.003

With such a close correspondence between physical and perceived distance between stimuli, we should expect a high correlation between decision-time and *d* as physically assessed. The results of this comparison are as follows: for D. C., tau = − .70; for V. C., − .55; for M. H., − .66; for P. K., − .69. For each tau *P* is less than .00001.

Experiment 2.—This experiment differed from Exp. 1 only with respect to the nature of stimuli. The results for the one *S*, P. K., may be sufficiently indicated by presenting the low, high, and median tau values, which were − .40, − .57, and − .47, with respective *P* values of .02, .001, and .007.

DISCUSSION AND CONCLUSIONS

The results of both experiments clearly sustain the hypothesis that decision-time and psychological distance (*d*) between the alternatives of a choice are negatively correlated. Implications of these results for measurement theory in particular and psychological theory in general are discussed below.

In 1902 Cattell argued against the feasibility of what is now called in Coombs' measurement theory the Method of Cartwheels. He did not think that *S* could with any certainty compare distances between stimuli. The consistency of choices in Exp. 1 and 2, as well as the close correspondence between the psychological and physical assessments of distance in Exp. 1, provide evidence contradictory to Cattell's assertion.

The stimuli of Exp. 1 were physically assessed. The physical measures agreed closely with the distance relations inferred from *S*'s choices. They could also be used to predict decision-time. Apparently for dimensions such as brightness, psychological relations are in close correspondence with physical. This by no means is to argue that physical measurement in general can substitute for psychological measurement. The results of Exp. 2, where the stimuli were words, emphasize the necessity of psychological measurement.

Decision-time has been shown to be consistent with the information contained in choices. Now this would remain theoretically significant but practically unimportant if decision-time provided no new information, since it is a much simpler procedure to collect choices than decision-times.

It is with respect to the *incomparable ds* that decision-time might be useful. Thus, having been validated against the comparable *ds*, decision-time can be used to order those *ds* which on the basis of the information in the choices are otherwise incomparable. As an example of this the data for subject M.H. may be examined. It is impossible to order the two *ds*, $\overline{AB}$-$\overline{BC}$ and $\overline{BD}$-$\overline{AB}$ from his choices. From the decision-times in Table 1, however, it can be inferred that $\overline{AB}$-$\overline{BC}$ is smaller than $\overline{BD}$-$\overline{AB}$, since the decision-time for the former is 16.93 sec. and that for the latter is only 3.54 sec.

The impossibility of obtaining from the choices a *simple* order on the *d*s, while an inconvenience in the testing of the hypothesized relation between decision-time and *d*, at the same time provides justification for the measurement and use of decision-time.

There are experiments (**1, 8, 9**) which have demonstrated a negative correlation between decision-time and response-probability. It might be expected that there would be no systematic differences in decision-time at a given level of response-probability. The results of the present experiments, of course, have proven this hypothesis false. The responses in these experiments were highly predictable and consistent, but the decision-times were nevertheless related as expected to distance between rims.

In these experiments each of the decision-time values was obtained from a *single* observation. Despite the numerous potential sources of unreliability which might influence such a delicate variable as decision-time, it was not found necessary to obtain an average over replications.

This fact has several implications. In the first place it is certainly more efficient to be able to rely on single observations than on many. Secondly, and more important, it is often desirable *not* to replicate, because of possible experimental dependencies between replications. Related to this, but on a more general level, is the point that averaging over replications on a single *S* may be no more defensible than averaging over *S*s. The results of these experiments show that single observations, both of decisions and decision-times, if properly obtained, can be meaningful.

Summary

The Method of Cartwheels, as recently developed by Coombs for the measurement of stimulus similarity, was applied to the problem of predicting decision-time. In Exp. 1, the stimuli were five shades of gray; in Exp. 2, five words. In both, 40 combinations (cartwheels) of four of the five stimuli were presented; the four stimuli of a cartwheel were arranged as two pairs (rims), and *S*'s task was to select the pair whose members were the more alike.

From *S*'s choices the stimuli and the distances between them could be ordered for him by means of the Unfolding Technique. A method was developed for using this information to obtain a measure of the difference (*d*) between the rims of each cartwheel. The decision-time associated with each cartwheel was also measured. As predicted, the two measures were negatively correlated for each of the 4 *S*s in Exp. 1, with median taus ranging from −.56 to −.73, and for the *S* of Exp. 2, with a median tau of −.47.

In Exp. 1 both the choices and decision-times were highly predictable from a physical assessment of similarity.

The correlations between decision-time and *d* were obtained under conditions of high choice consistency. The decision-time data were based on single observations.

References

1. Birch, D. Discrimination learning as a function of the ratio of nonreinforced to reinforced trials. *J. comp. physiol. Psychol.*, 1955, **48**, 371–374.
2. Cattell, J. McK. The time of perception as a measure of differences in intensity. *Phil. Stud.*, 1902, **19**, 63–68.
3. Coombs, C. H. Psychological scaling without a unit of measurement. *Psychol. Rev.*, 1950, **57**, 145–158.
4. Coombs, C. H. A theory of psychological scaling. *Univer. Michigan. Engng. Res. Inst. Bull.*, 1951. No. 34. Pp. vi, 94.
5. Coombs, C. H. Theory and methods in social measurement. In L. Festinger & D. Katz (Eds.), *Research methods in the behavioral sciences*. New York: Dryden, 1953.
6. Coombs, C. H. A method for the study of interstimulus similarity. *Psychometrika*, 1954, **19**, 183–194.
7. Dember, W. N. Decision-time and psychological distance. Unpublished doctor's dissertation, Univer. of Michigan, 1955.
8. Kellogg, W. N. The time of judgement in psychometric measures. *Amer. J. Psychol.*, 1931, **43**, 65–86.
9. Rosenblith, W. A., & Stevens, K. N. On the DL for frequency. *J. acoust. Soc. Amer.*, 1953, **25**, 980–985.
10. Spence, K. W. The relation of response latency and speed to the intervening variables and *N* in S-R theory. *Psychol. Rev.*, 1954, **61**, 209–216.

(Received December 13, 1955)